Quantitative Genetics in Maize Breeding
HANDBOOK OF PLANT BREEDING

Editors-in-Chief:

JAIME PROHENS, Universidad Politécnica de Valencia, Valencia, Spain
FERNANDO NUEZ, Universidad Politécnica de Valencia, Valencia, Spain
MARCELO J. CARENA, North Dakota State University, Fargo, ND, USA

Volume 1
Vegetables I: Asteraceae, Brassicaceae, Chenopodicaeae, and Cucurbitaceae
Edited by Jaime Prohens and Fernando Nuez

Volume 2
Vegetables II: Fabaceae, Liliaceae, Solanaceae and Umbelliferae
Edited by Jaime Prohens and Fernando Nuez

Volume 3
Cereals
Edited by Marcelo J. Carena

Volume 4
Oil Crops
Edited by Johann Vollmann and Istvan Rajcan

Volume 5
Fodder Crops and Amenity Grasses
Edited by Beat Boller, Ulrich K. Posselt, Fabio Veronesi

Volume 6
Quantitative Genetics in Maize Breeding
Arnel R. Hallauer, Marcelo J. Carena, J.B. Miranda Filho
To my wife Jan, daughter Beth, and son Paul
   – Arnel R. Hallauer

To Irene and our ‘hybrids’ (Martin y Diego)
To my parents in Argentina (abuelos Bob y Cris)
   – Marcelo J. Carena
Preface

Plant breeding is a science of evolution. The scientific basis of plant breeding started in the 1900s. The rediscovery of Mendelian genetics and the development of the statistical concepts of randomization and replication had considerable impact on plant breeding methods. They provided a genetic basis for the variation observed among individuals, separating genetic and environmental effects, and valid experimental techniques for measuring those differences. Breeders work with traits and environments and their major task is to increase the frequency of favorable alleles of quantitative traits, controlled by a large, unknown number of genes interacting with target environments.

Maize is an economically important crop for feed, fiber, fuel, and food. It is used as an ingredient in an endless list of manufactured products that affect the nutrition of the world’s population. Maize is a cross-pollinated species with unique and separate male (tassel) and female (ear) organs. Maize breeding has unique features that are different from the other extensively cultivated grain species which are primarily self-pollinated. Techniques from both self- and cross-pollinated crops are utilized in maize that allow production of large sample sizes easily and a diversity of breeding methods. These methods are geared toward the development of improved populations, inbred lines, and their hybrids for different types of markets.

Maize breeding is one of the successful examples of breeder-directed evolution. Breeders have been effective in developing improved cultivars to meet the changing cultural and environmental conditions of the past 150 years. Modern maize breeding methods are primarily a 20th century phenomenon. Applied maize breeding has been effective in developing improved hybrids during the past 100 years. The inbred–hybrid concept was developed in the public sector and is still considered one of the greatest achievements in crop breeding. It is considered one of the most successful plant breeding breakthroughs as a consequence of private–public research cooperation and business vision. Development of the commercial seed industry is testimony of breeding methods that have evolved for the economical production of high-quality hybrid seed that is accepted and demanded by the modern farmer. In order to meet the increasing and divergent uses of maize, breeding methods have evolved to increase the effectiveness and efficiency of selection for several quantitative traits in multiple stages and environments often at a rate of more than one season per year due to winter nurseries.
Maize breeders decide which combination of traits and environments is needed to breed for both inbreds and hybrids. A trait controlled by genes that are not significantly affected by the environment can be improved very effectively. For instance, breeders have the option to improve flowering date for earliness at a rate of 2–3 days per year at a minimum cost by stratified mass selection (e.g., less than a penny per plant screened) or utilized cloned genes (e.g., vgt1) for a more expensive procedure (e.g., if a grant is awarded). Scientists have also decided to invest federal funds in understanding the genetic basis of flowering time in a specific maize population (e.g., NAM) which should alert us of the need to increase maize sampling while linking classical and modern quantitative genetic approaches. However, most traits that concern the breeder (e.g., yield, drought tolerance) are economically very important, difficult to measure, and quantitative in nature often with significant environmental influence on the trait’s expression. For instance, breeding approaches for drought tolerance exploiting polygenic effects with possible transgenes would seem desirable. Though single-gene transgenic approaches are currently being proposed by industry, choosing the right methodology (e.g., may be an interaction of both) will demonstrate there is no limit to genetic improvement for drought tolerance when most tolerance genes are targeted in the breeding process and not just a single-gene approach. Drought-tolerance efforts on polygenic effects do not have a limit on genetic improvement and quantitative traits, at the moment, are better explained by polygenes rather than quantitative trait loci (QTL).

Quantitative traits are controlled by a large, but unknown, number of genes, each having a small effect on the total expression of the trait. The environment in which they are measured determines their effects and it is dependent on genetic background as each hybrid has its own genetic effects. These traits are characterized by degree differences among phenotypes that do not fall into distinct categories. They have a complex inheritance including dominance, epistasis, linkage, and the interaction of genetic and environmental effects. Quantitative traits are controlled by the joint action of many genes, and genetic improvement of plants has been successful even though the knowledge of genes controlling these traits was minimal.

Unlike other crops maize is a model crop for achieving significant genetic gain with large genetic variability. Genetically broad-based public germplasm has significantly been utilized and recycled by industry before intellectual property rights were available. B14, B37, B73, and B84 are examples of publicly developed inbred lines that have generated billions of dollars to the agricultural sector via their use in hybrids. They were derived from a breeding scheme that integrated germplasm improvement (e.g., five cycles of half-sib intra-population recurrent selection before developing B73) with inbred line development. The development of these lines needed continuous federal and state funding through decades for creating the genetically broad-based synthetic variety Iowa Stiff Stalk Synthetic (BSSS), improving it by five cycles of half-sib recurrent selection with a tester, conducting several years of inbreeding and extensive hybrid testing, and producing seed increases for release and industry distribution. Few of these programs are left due to public funding restrictions and lack of commitment from breeders to conduct long-term
programs for germplasm improvement. Moreover, such programs are often indirectly discouraged due to land-grant University tenure and royalty concerns as well as encouragement for short-term grants with indirect costs often leading scientists to avoid unbiased decisions and research. But, scientists are also encouraged to keep long-term breeding goals active as they provide an extensive and annual production of peer-refereed high-quality manuscripts with applied science impact.

Although the basic breeding methods for development of maize inbred lines and hybrids were described by 1910, significant contributions have been visualized and tested for modernizing the basic breeding methods including transgenic maize for single genes. However, as for the drought example, there is no limit to genetic improvement when most genes are targeted in the breeding process. Even though choice of base germplasm continues to be the priority we still have the challenge to integrate effective molecular tools for selecting quantitative traits that are either difficult or expensive to measure with low heritability (e.g., root traits, fast dry down). Definitively, modern technology should target the right traits, those currently challenging to phenotype. The genome of B73 has been sequenced. The application of this discovery, however, still is not fully realized in breeding programs. Therefore, there might be alternative and/or complementary options to speed up the development of maize hybrids. We definitively need the integration of basic and applied science for common long-term goals directed at better cultivars. Heterotic effects are unique for each hybrid and sequencing efforts on only B73 and others alike might limit the identification of useful and unique alleles (e.g., tropical and/or early maturing genetic backgrounds) for complex traits, ultimately the desirable ones to broaden our germplasm base.

Similar to other scientific disciplines, many changes have occurred in the methods, techniques, information available, and germplasm used in maize breeding since the first edition of this volume. Interest in the inheritance of quantitative traits of maize was one of the basic areas of maize research during the 30-year span of 1945–1975. During this time frame, there were areas of maize breeding of concern because of the seemingly yield plateau of double-cross hybrids; the germplasm available for breeding purposes and how they could be improved; selection and breeding methods to enhance effectiveness of developing inbred lines and hybrids; and the one area of common interest in all facets of maize breeding – the genetic basis of heterosis. Extensive data have been reported for the types of genetic effects important in quantitative trait expression across different types of maize populations and hybrids. However, against predictions and multi-million dollar grants a uniform theory on heterosis remains inconclusive. It is generally agreed, however, that the presence of non-additive genetic effects (dominance and epistasis) are necessary even though measuring epistasis continues to be challenging to also predict at the molecular level. Because non-additive effects seem so important for heterosis expression, it seems that relative importance of the different non-additive genetic effects will vary among hybrids because of the combination of genes and alleles that is unique for each highly productive hybrid.

Our main concern in this volume is to describe how the principles of quantitative genetics and cyclical selection schemes have been used in maize breeding
research. In addition to the principles discussed, data are summarized and updated from reported studies. It is not intended that the breeding methods discussed in this volume will, or should, replace those currently used in the public and private sectors; they are intended to supplement those currently used and help devise new ones. The integration of the breeding methods discussed herein with those currently used should enhance future breeding efforts to maintain the level of genetic gain of the past 70 years and help achieve industry targets (e.g., double genetic gains in 30 years). Future genetic gains in maize hybrid yields are dependent on the incorporation of useful and unique genetic diversity. The integration of breeding methods discussed in this volume should enhance future breeding efforts to maintain and create active public breeding programs, not only adapting for improving genetically broad-based germplasm, but also developing products and training the next generation of maize breeders. There are fewer applied public maize breeding programs than generally thought. It will require the combine support of industry, grower associations, grant agencies, and land-grant systems to help keep applied public maize breeding programs strong.

One aspect of genetics that has dramatically changed how maize breeders design and conduct breeding programs is the vast and rapid developments in molecular genetics and the generation of genotypic data at a lower cost. Similar to the study of quantitative genetics for the types of genetic variation and inheritance of maize traits, certain techniques of molecular genetics also have been extensively studied and applied to maize breeding programs (e.g., marker-assisted backcrossing for single-gene trait integration). Quantitative genetics and molecular genetics represent the opposites. Data for the quantitative genetic studies were collected on the phenotypes of families and progenies replicated within and across environments. In contrast, the molecular geneticists study gene and allele effects at the DNA level. The disparity between the two facets of genetics may seem irreconcilable, but they are becoming more interwoven with continued advancements in molecular genetics and the need for ‘phenotyping.’ Interest and emphasis on the study on the inheritance of quantitative traits decreased rapidly with the advent and rapid expansion of molecular genetics. However, they seem to be different approaches for a similar information generation. Initially, molecular genetics had the desire to study complex traits but emphasized the study of traits with major effects, such as resistance and/or tolerance to major pests (insects, diseases, and weeds) that affect maize production. If molecular genetics was to have greater impact it became obvious that it had to examine how molecular genetic techniques could be used to improve the important economic traits, such as grain yield. The identification and use of molecular markers to assist in selection of quantitative trait loci (QTLs), genome-wide selection, and association mapping have become common areas of research. Breeders are trying to sort out what can be applied. The integration of quantitative and molecular genetic information will only increase in the future as sample sizes approach those of the past quantitative genetic studies. Cooperation between the public and private sectors for latest technology seems to be a solution to avoid spending precious resources with state and federal funds in labs that become obsolete quickly. It also allows the public scientist to generate basic research as a matter of weeks
vs. years. The maize genome has over 60,000 genes (number increases annually). Therefore, all aspects of genetics will be needed for continued genetic advances. The integration of molecular biologists, geneticists, physiologists, and breeders is key. Maize breeders will need all the resources available to them to manipulate the large number of genetic factors that affect the traits important in developing elite inbred lines to produce genetically superior hybrids with consistent performance in changing environments. Exploiting unique environments (e.g., North Dakota) can accelerate desirable evolution toward adaptation to climate change.

The objective of this volume, as part of the *Handbook of Plant Breeding*, is to increase awareness of the relative value and impact of applied maize breeding for all crop uses and its security and sustainability. The topics included should be of interest of graduate students in plant breeding and to breeders conducting research on not only breeding and selection methods but also developing pure lines and hybrid cultivars in maize and other crop species. Because of rapid development of the inbred line–hybrid concept in maize, chapters pertain mostly to maize but are applicable to other crop species. The volume’s main contribution will be to breeders, geneticists, students, and policy makers willing to work together for the long-term sustainable crop improvement and production for various environments. We hope long-term genetic maize improvement programs with genetically broad-based genetic materials (e.g., GEM/ EarlyGEM programs) are encouraged.

The authors wish to acknowledge the contributions and suggestions others have provided in the preparation of this volume. This project was initiated to commemorate 20 years of the last edition of this book, a very useful resource for research, teaching, and breeding. Our target was to update content and make a unique style presentation. Corrections of errors detected by the authors and others have been made, some changes were made in notation for consistency, rewriting was done for clarity, and content and references on the subject matter were added.

The authors owe a deep gratitude to Mary Lents for all of her efforts for the past 30 years, which made this book possible. Her diligence, talents, interests, persistence, attention to details, typing skills, and editing link ‘between Iowa and North Dakota’ are sincerely appreciated. She accomplished these tasks in addition to her regular duties as secretary. We sincerely appreciate all her past contributions and wish her only the best in the future, especially at this moment of uncertainty.

This would not have been possible without the initiative and exchange of ideas with Hannah Schorr and Jaime Prohens.

We also want to appreciate the support of the NDSU and ISU maize breeding units and the North Dakota Corn Growers Association and North Dakota Corn Council Utilization, especially Duane Wanner and Paul White for helping generate much of the data in these units. Also, a special thanks for the comments, reviews, and ideas of Junyun Yang, Santosh Sharma, and Tonette Laude, current graduate students at NDSU.
Special thanks to Irene Santiago and Jan Hallauer during the past two years of work on this volume and especially their continuous support.

Ames, IA, USA
Fargo, ND, USA

Arnel R. Hallauer
Marcelo J. Carena
# Contents

1 Introduction ................................................................. 1  
   1.1 Quantitative Genetics ................................................... 9  
   1.2 Population Improvement: What Do We Mean 
       by Recurrent Selection? .............................................. 11  
   1.3 Inbred Line Development ............................................. 16  
   1.4 Conclusions ............................................................. 25  
   References ................................................................ 28  

2 Means and Variances ......................................................... 33  
   2.1 Genetically Narrow- vs. Broad-Based Reference Populations .... 33  
   2.2 Hardy–Weinberg Equilibrium .......................................... 34  
   2.3 Means of Non-inbred Populations and Derived Families ....... 36  
   2.4 Means of Inbred Populations and Derived Families ............ 41  
   2.5 Mean of a Cross Between Two Populations ....................... 42  
   2.6 Average Effect ............................................................ 43  
   2.7 Breeding Value ............................................................ 45  
   2.8 Genetic Variance ......................................................... 46  
   2.9 Means and Variances in Backcross Populations ................. 56  
   2.10 Heritability, Genetic Gain, and Usefulness Concepts ......... 57  
   2.11 Generation Mean Analysis ............................................ 60  
   References ................................................................ 66  

3 Resemblance Between Relatives ............................................. 69  
   3.1 Introduction ................................................................. 69  
   3.2 Theoretical Basis of Covariance ...................................... 70  
   3.3 Covariance Between Relatives as a Linear Function 
       of Genetic Variances .................................................... 71  
   References ................................................................ 78  

4 Hereditary Variance: Mating Designs ...................................... 81  
   4.1 Bi-parental Progenies ..................................................... 86  
   4.2 Pure Line Progenies (Analysis in Self-Pollinated Crops) ....... 90  
   4.3 Parent–Offspring Regressions ......................................... 91  
   4.4 Design I ................................................................. 96  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.6 General vs. Specific Combining Ability</td>
<td>415</td>
</tr>
<tr>
<td>References</td>
<td>418</td>
</tr>
<tr>
<td>9 Inbreeding</td>
<td>425</td>
</tr>
<tr>
<td>9.1 The Need for Maize Artificial Pollination</td>
<td>425</td>
</tr>
<tr>
<td>9.2 Early Reports of Inbreeding</td>
<td>428</td>
</tr>
<tr>
<td>9.3 Inbreeding Systems</td>
<td>431</td>
</tr>
<tr>
<td>9.4 Inbreeding Due to Small Population Size</td>
<td>436</td>
</tr>
<tr>
<td>9.5 Estimates of Inbreeding Depression</td>
<td>442</td>
</tr>
<tr>
<td>9.6 Frequency of Useful Lines</td>
<td>456</td>
</tr>
<tr>
<td>9.7 Types of Hybrids Produced from Inbred-Lines</td>
<td>458</td>
</tr>
<tr>
<td>9.8 Heterozygosity and Performance</td>
<td>464</td>
</tr>
<tr>
<td>References</td>
<td>471</td>
</tr>
<tr>
<td>10 Heterosis</td>
<td>477</td>
</tr>
<tr>
<td>10.1 Introduction and Major Achievements</td>
<td>477</td>
</tr>
<tr>
<td>10.2 Empirical Evidence</td>
<td>480</td>
</tr>
<tr>
<td>10.3 Genetic Basis</td>
<td>493</td>
</tr>
<tr>
<td>10.4 Biometrical Concept</td>
<td>497</td>
</tr>
<tr>
<td>10.5 Heterosis and Prediction Methods Across Genotypes</td>
<td>499</td>
</tr>
<tr>
<td>10.6 Components of Heterosis in Inter-varietal Diallel Crosses</td>
<td>515</td>
</tr>
<tr>
<td>10.7 Conclusions</td>
<td>519</td>
</tr>
<tr>
<td>References</td>
<td>523</td>
</tr>
<tr>
<td>11 Germplasm</td>
<td>531</td>
</tr>
<tr>
<td>11.1 Origin of Maize</td>
<td>534</td>
</tr>
<tr>
<td>11.2 Classification of Maize Germplasm</td>
<td>538</td>
</tr>
<tr>
<td>11.3 Races of Maize in the Western Hemisphere</td>
<td>542</td>
</tr>
<tr>
<td>11.4 European Races of Maize</td>
<td>545</td>
</tr>
<tr>
<td>11.5 US Corn Belt Germplasm</td>
<td>548</td>
</tr>
<tr>
<td>11.6 Germplasm Improvement</td>
<td>552</td>
</tr>
<tr>
<td>11.7 Potential and Use of Exotic Germplasm</td>
<td>558</td>
</tr>
<tr>
<td>References</td>
<td>570</td>
</tr>
<tr>
<td>12 Breeding Plans</td>
<td>577</td>
</tr>
<tr>
<td>12.1 Choice of Germplasm</td>
<td>581</td>
</tr>
<tr>
<td>12.2 Recurrent Selection and Germplasm Improvement</td>
<td>582</td>
</tr>
<tr>
<td>12.3 Integrating Recurrent Selection with Cultivar Development</td>
<td>591</td>
</tr>
<tr>
<td>12.4 Intra-Population Genetic Improvement</td>
<td>592</td>
</tr>
<tr>
<td>12.5 Inter-Population Genetic Improvement</td>
<td>609</td>
</tr>
<tr>
<td>12.6 Additional Considerations for Germplasm Improvement</td>
<td>631</td>
</tr>
<tr>
<td>12.7 Additional Considerations for Inbred Line Development</td>
<td>647</td>
</tr>
<tr>
<td>References</td>
<td>649</td>
</tr>
<tr>
<td>Index</td>
<td>655</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

Maize (Zea mays L.)-directed evolution through breeding started when humans realized the potential of the species for food, feed, fiber, and fuel. Although the ancestral pedigree of maize has not been fully resolved yet, the early maize breeders certainly played an important role in domesticating and developing the species as we know it today. Maize, however, is known to be one of the few major cultivated species indigenous to the Western Hemisphere about 7,000–10,000 years ago (Wilkes, 2004).

The transition of maize from a weed species to a cultivated crop species required years and patience of the early maize breeders. Depending on the preferences and environments of the different Native American settlements, different selection pressures for various traits were applied. Growers could effectively select and fix traits such as kernel color and texture for a range of environments. Because all harvesting was done on an individual plant basis, variation among plants and ears for easily recognizable traits (e.g., earliness) would be noted. Nature, on the other hand, would play an important role in the development of maize strains that were resistant to pests; day-length sensitivity; and drought, heat, and cold tolerance. The traits that people could select and fix were transferable from one local area to another; people and nature would cooperatively develop maize landraces adapted for wide differences in environments. Local settlements probably used certain traits as genetic markers for their particular populations. Accidental mixtures and cross-pollination would contribute to the wide variation of maize varieties that was present before Columbus arrived in the Western Hemisphere. For instance, more than 300 years ago the women of agricultural tribes of Native Americans (Mandan, Arikara, and Hidatsa) living in the Upper Missouri Valley were the first ‘corn breeders’ in the northern USA (Olson and Walster, 1932). Maize was the first agricultural crop in North Dakota (ND) and Lewis and Clark were able to survive the winter of 1804–1805 due in part to maize supplied by Native Americans. These early plant breeders took care of maize for its survival. They referred it as the ‘corn mother,’ ‘the woman that never dies,’ which resulted in diverse strains of short-season maize varieties. These adapted varieties were one of the sources of short-season maize germplasm available when the North Dakota Agricultural College was founded in 1890.
Selection procedures used by early maize breeders would seem primitive compared with present-day breeding methods, but they recognized the traits needed to sustain their civilizations. The effectiveness of selection by early maize breeders is evident from the hundreds of races and thousands of varieties that have been collected and described. Mass selection, defined as ‘simple selection’ in ND (Olson et al., 1927), was performed by selecting ears after harvest (method I) and, even at that time, selecting ears from the best plants in the field (method II). The latter tried to answer the same questions asked by Gardner (1961) when trying to reduce environmental bias selecting plants with stiff stalks and healthy leaves before looking at the ear. As a consequence of these simple selection procedures maize can be grown throughout almost the complete range of altitudes and latitudes around the world. Certainly the greatest maize breeding advance was made by converting a wild weed species of relatively low yield to a cultivated crop species that has yielded over 15 t/ha (tons per hectare) under modern cultural practices. On the other hand, this progress has been at the expense of genetic diversity by utilizing very few landraces within one race.

Maize breeding has evolved through several stages. Because maize was indigenous to the Western Hemisphere, European settlers soon were introduced to maize and depended on it for survival. As the European migration expanded along the coastlines and into the interiors of the Western Hemisphere, pioneers would have come into contact with a great diversity of maize germplasm. Movement of the people and exchange of maize germplasm permitted the introgression of divergent varieties and complexes of the Western Hemisphere. In the Northern Hemisphere, expansion of the cultivation of the Northern Flint complex and the Southern Dent complex led to the development of the highly productive US Corn Belt Dent varieties (Brown, 1950; Anderson and Brown, 1952). The Northern Flints and Southern Dents were distinct complexes, but the expansion in maize cultivation southward (for the Northern Flints) and northward (for the Southern Dents) resulted in a reciprocal introgression of the two complexes (Hudson, 2004). Seeds of the two distinct landrace complexes were carried by people, but the crossing between the two complexes probably occurred more by contamination than by planned crosses. Crossing of the two distinct complexes created a vast reservoir of genetic variability for plant and ear traits. Simple mass selection based on individual plants, therefore, was an effective breeding methodology for developing varieties that possessed traits appealing to growers and early colonial maize breeders. The hybrid swarms arising from crossing the Northern Flint and Southern Dent complexes would not have been as extensive as in pre-Columbian times, but the range in variability was great enough for the development of varieties having distinctive plant and ear traits.

Development of the US Corn Belt Dent varieties was not a planned breeding program. As the interior of the USA was developed, seeds of settlers’ varieties were brought westward from the Atlantic seacoast. Each individual seed lot would have been subjected to different selection pressures, depending on the individuals selecting seed to propagate the crop the following season. In some instances, careful selection was given to ear traits (row number, flint vs. dent, color, prolificacy, etc.), whereas others selected for early maturity, shorter plants, freedom from tillers, and plant type. Selection, as in pre-Columbian times, developed varieties with distinctive
traits and adapted to specific environments. As cultivation of maize became more extensive, introgression among the selected strains occurred because isolation was reduced. The amount of introgression among selected strains depended on the topography of the areas and the amount of interchange of seed among pioneer settlers. Genetic divergence among selected strains was sufficiently great that crosses among them suggested that variety hybrids were superior to growing the selected varieties per se. Controlled crosses between selected varieties improved grain yield performance.

Beal (1880) suggested the potential of variety hybrids from seed produced from controlled crossing of two varieties; although variety hybrids were superior to the parental varieties, they apparently were never used to any great extent. However, the potential of the population-hybrid concept for certain improved varieties is worth investigating economically for certain agriculture sectors that rely on low-cost seed production alternatives (Carena, 2005a). A brief summary of the early history of maize breeding was given by Goodman (1976).

Two US developments in the early part of the 20th century had a profound effect on modern maize breeding: (1) development of the maize show card for exhibiting maize ears in national maize shows had an effect on selection procedures of farmers, producers, and breeders for selecting ear types that conformed to show card standards and (2) public research reported independently by Shull (1908, 1909) and East (1908) outlined the basis of the breeding methods for developing and producing modern maize hybrids. The development of the maize show card had only a temporary effect on maize breeding, whereas public hybrid research formed the foundation of modern maize breeding and its successful profitable business.

Development of the maize show card created further interest in selection within maize varieties because intense selection pressure was given to ears that conformed to the standards of the show cards (Table 1.1).

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>US maize growers and the score card utilized in ‘corn shows’ (Bowman and Crossley 1908)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traits</td>
<td>Points</td>
</tr>
<tr>
<td>I. General appearance</td>
<td>25</td>
</tr>
<tr>
<td>1. Size and shape of ear</td>
<td>10</td>
</tr>
<tr>
<td>2. Filling of butts and tips</td>
<td>5</td>
</tr>
<tr>
<td>3. Straightness of rows</td>
<td>5</td>
</tr>
<tr>
<td>4. Uniformity of kernels</td>
<td>5</td>
</tr>
<tr>
<td>II. Productiveness</td>
<td>60</td>
</tr>
<tr>
<td>1. Maturity</td>
<td>25</td>
</tr>
<tr>
<td>2. Vitality</td>
<td>25</td>
</tr>
<tr>
<td>3. Shelling percentage</td>
<td>10</td>
</tr>
<tr>
<td>III. Breed type</td>
<td>15</td>
</tr>
<tr>
<td>1. Size and shape of ear</td>
<td>5</td>
</tr>
<tr>
<td>2. Size, shape, and dent of kernel</td>
<td>5</td>
</tr>
<tr>
<td>3. Color of grain</td>
<td>2</td>
</tr>
<tr>
<td>4. Color of cob</td>
<td>2</td>
</tr>
<tr>
<td>5. Arrangement of rows</td>
<td>1</td>
</tr>
</tbody>
</table>
Individual ear and small-ear sample winners of shows often commanded high premiums when sold as a seed source. Although fields of maize were open pollinated, bottlenecks for genetic variability often would have resulted from the small number of ears representing the winning samples of maize shows. Usually the winning samples resulted from the time and patience the exhibitor spent inspecting fields for ears that were uniform in appearance and conformed to show card standards. Time and patience in selecting ear types resulted in a number of named varieties that had distinctive traits. Often several sub-strains of a limited sample that had won top honors at a prestigious maize show were developed. Although maize shows were very popular in the early part of the 20th century, critical comparisons of prize-winning samples that met score card standards with samples that deviated from them showed that esthetic ear traits under above average environmental conditions were not an indicator of yield performance. Results of comparative yield tests quickly led to the demise of the maize shows when it was demonstrated that a ‘Krug Yellow Dent’ variety that was not selected to meet score card standards was equal to a ‘Reid Yellow Dent’ variety that was. This was accentuated by the current dilemma between yields per se and yield per acre.

Probably the most well-known public scientist in early hybrid corn research is G.H. Shull whose maize research records date back to 1904. At that time research focused on heredity as a basis for improving plants and animals. Shull studied the theory of genetics and its application to plant breeding. The research of East (1908), Shull (1908, 1909, 1910), and Jones (1918), contrary to the development of show score cards as selection criteria, provided the framework for maize breeding that is still used today. Although the essential features of modern hybrid maize breeding were outlined before 1920, several distinct phases have been recorded across decades (Fig. 1.1). Single-gene trait integration for one-, three-, and currently eight-event hybrids (e.g., system proposed by Monsanto and Dow Agrosciences) is an

Fig. 1.1  Sequence of breeding methods developed for production of hybrids in the USA
additional challenging phase. It requires a detailed system of conversions, testing, integration with inbred line development, and data management.

Shull (1908) and East (1908) showed that an open-pollinated maize variety consisted of a range of heterogeneous genotypes from which homozygous and homogeneous pure lines (inbred lines) could be developed. Controlled pollinations of pollen (male gametes) on silks (female gametes) of the same plant during five to seven generations of selfing produced inbred lines. Although inbred lines that were uniform and repeatable for phenotype were obtained by self-pollination, the inbred lines often were weak and difficult to propagate. However, when certain of the weak inbred lines were crossed, vigor was restored and the yield of the inbred line crosses (or single-cross hybrids) usually exceeded the one from the original open-pollinated variety from which the lines were developed. The single crosses were uniform (homogeneous) because they were genetically identical heterozygous and homogeneous plants. These observations of inbreeding and crossing led Shull (1910) to describe a method for producing maize hybrids. His remarkable conclusions gave a lucid outline of the subsequent course of maize breeding that were based on the limited data and experience available at that time. Remarkable too was Shull’s conclusion that the inbred–hybrid concept had no practical value due to small amount of seed produced on weak inbred lines. This led Shull to discontinue his research efforts by 1916.

East at Connecticut and Shull at Cold Spring Harbor independently started studies of inbreeding and crossbreeding in 1906 (Hayes, 1963) and provided essential insight into the efforts of maize inbreeding (East, 1908; Shull, 1909). Because of the poor vigor and seed production of the inbred lines, East was also somewhat pessimistic that the procedure would have practical usefulness. East produced the first Leaming inbred lines and was directly influenced by the biological principles of Darwin, Mendel, Festetics, and Vilmorin in relation to plant improvement (Hayes, 1956). East related those principles to the more practical plant improvement studies to achieve his goals. It was not until Jones (1918) suggested that two single crosses be used as parents in the production of hybrids that it seemed hybrid maize would become a reality. Jones suggested to East a procedure that would make hybrid maize a reality for industry and farmers, using the already developed Leaming lines as females and Burr’s White inbreds as males (Jones, 1918). There were concerns that combining four inbred lines to make double-cross hybrids would be a step back to open-pollinated maize and that performance would be reduced compared to single-cross hybrids. However, producing seed on a single-cross parent provided a solution to marketing hybrids, annually to farmers. The change from open-pollinated varieties to double-cross hybrids was a significant improvement in developing maize hybrids with improved standability and grain yield performance. As a consequence, breeding efforts toward population improvement and genetic diversity were reduced.

The potential of seed production on double-cross hybrids stimulated research to develop inbred lines during the 1920s. After years of intensive and at times discouraging research searching for effective inbred lines, the largest maize research organization in the private sector, ‘Funk Farms,’ produced the first maize double-cross
hybrid (US13) using lines developed in the public sector (Crabb, 1947). Also, the first Pioneer hybrid for sale was ‘Copper Cross’ in 1924 and included East’s Leaming lines on the female side (Hayes, 1963). These lines were developed from the US Corn Belt Dent variety Leaming (Leaming, 1883; Lloyd, 1911).

It soon became evident that it was simpler to develop inbred lines than it was to determine their worth in hybrids especially for quantitative traits. For instance, it was recognized, contrary to the expected, that cold intolerant lines sometimes gave outstanding short-season hybrids for ND. Not all lines in hybrid combinations were superior to the original open-pollinated varieties. The variation among double-cross hybrids showed that only certain combinations of inbred lines gave a satisfactory double-cross hybrid which was an early challenge maize breeders were facing. Preliminary screening techniques by Davis (1927) and Jenkins and Brunson (1932) showed that crossing lines with a common tester (procedure often referred to either topcross or testcross by maize breeders) was effective for discarding lines that did not have satisfactory performance in hybrids. Many lines could be discarded on the preliminary testcross performance information. Often a sizable number of lines that had potential for use in hybrids still were available. Even if only 20 lines were available for testing in hybrids, 190 different single crosses and 14,535 different double crosses were possible. As a consequence, Jenkins (1934) developed and tested four methods for predicting double-cross hybrids from use of single-cross hybrid data. Experimental data showed the reliability of method B (average performance of four non-parental single crosses) for reducing the amount of testing needed to determine better performing double-cross hybrids. Testcross tests for preliminary screening of lines and prediction methods for reducing the number of double-cross hybrids for testing were widely used and contributed greatly to the efficiency of maize breeding at that time.

Seven generations of selfing are required to obtain lines that are nearly homozygous and homogeneous. Testcross trials and prediction methods aided testing of lines in hybrids after they were inbred to near homozygosity. The next logical step was to establish whether the hybrid yielding ability of lines could be determined before they approached homozygosity. Selection during inbreeding can be effective for some traits (e.g., plant type, resistance to attack by pests, kernel color, and male and female characteristics such as ability to shed pollen, ear size), but yield in hybrids was usually determined by producing crosses and testing when lines were nearly homozygous. If yield potential of lines in crosses could be determined before homozygosity, many lines could be discarded based on their combining ability to reduce the number continued by selfing. Jenkins (1935) and Sprague (1946) presented data showing that yield potential of lines in crosses could be determined in early generations of inbreeding. Consequently, lines that lacked yield contributions to crosses could be discarded in early generations of inbreeding; this permitted better sampling of populations for superior lines and the number of lines continued by inbreeding could be substantially reduced before they were homozygous. In addition, correlations between early- and late-generation testcrosses did show that the differences between performance of early- and late-generation testcrosses were primarily due to non-genetic factors (Bernardo, 1991).
Because of the development of maize breeding methodologies, formal breeding programs expanded rapidly in scope and number during the 1930s and 1940s. Hundreds of inbred lines were developed and tested. Cooperation and interchange of inbred lines among maize breeders identified elite inbred lines that were used extensively in the production and growing of double-cross hybrids. Acceptance of double-cross hybrids by farmers depended on the availability of seed and their knowledge of new seeds. Fig. 1.2 shows that nearly 100% of the maize acreage was planted to hybrids in Iowa by 1943 and in the USA by 1960. Hybrid production in ND was initiated in the 1930s. One decade later, the percentage of maize land planted to hybrids in ND was 12% ranging from 5% in northern and western ND (more marginal areas) to 35% in southeastern ND by 1941 (Wiidakas, 1942). This percentage increased to 50% by 1948 (80% in southeastern ND). Open-pollinated varieties were rapidly replaced by hybrids in the more favorable maize growing areas of the state. A very positive consequence of hybrid maize was the increase of productivity and stiff stalks under farm mechanization. Hybrids yielded 5–28% more than later and high-yielding varieties such as Minnesota 13. Seed costs, state and federal subsidies, and grower migration to urban areas also increased. Early-maturing hybrids were needed for the remainder of ND and popular open-pollinated varieties (e.g., Falconer) were still a safe choice for short-season areas. The first long-term North Dakota State University (NDSU) corn breeder, William Wiidakas, joined the ND Agricultural College in 1934 after earning an MS degree from University of Minnesota under the direction of H. K. Hayes. Therefore, NDSU maize breeding program probably had some influence of East’s group. Several

Fig. 1.2  Rate of acceptance of double-cross hybrids by farmers in the USA and Iowa
useful yellow-dent inbred lines (e.g., ND203 and ND230) were produced and successful early-maturing yellow-dent double-cross hybrids ('NODAK' hybrids) were released, produced by seed companies, and grown extensively in the state. These were early yellow-dent hybrids ranging from 75 to 90 RM (Wiidakas, 1942) and yielded up to 37% more than popular open-pollinated varieties.

Although double-cross hybrids had essentially replaced open-pollinated varieties by the 1950s, there was some concern that a yield plateau had been attained. Double-cross hybrids were used because it seemed the only feasible method of producing maize hybrids from the experiences of East (1908), Shull (1908, 1909), and Jones (1918). During the late 1950s and early 1960s a few single-cross hybrids were produced and grown. Improvement in management (use of herbicides, pesticides, and fertilizers) and in inbred lines (improved by selection, in many instances, recycled from earlier developed lines) permitted the production of single-cross hybrids that were economically feasible for producers and farmers. Double-cross hybrids require four parental inbred lines, whereas single-cross hybrids require two. Selection pressure, therefore, would be 100% greater among inbred lines for use in single-cross hybrids. Although elite inbred lines that were extensively tested were used for producing hybrids, trade-offs usually were necessary in the selection of lines to compensate for their known weaknesses; this would be easier for two than for four elite inbred lines. Cockerham (1961) also showed theoretically that expected genetic variability and expected genetic gain among single-cross hybrids within a population would be at least double that among double-cross hybrids. From field experience and theoretical considerations, the change from production and growing of double-cross hybrids to single-cross hybrids has increased in the USA from about 0% in 1960 to nearly 100% at the turn of the century. Because of the development of field husbandry and maize breeding methodology, the circuit was completed from Shull’s first suggestion of single-cross hybrids in 1908 to the present (Fig. 1.1).

The extensive use of recycled lines that followed the inbred–hybrid concept led to the introduction, adaptation, evaluation, and improvement of exotic germplasm resources for use in breeding programs. These pre-breeding efforts were initiated in the public sector by individuals who appreciate the possible contributions of exotic germplasm in order to increase the genetic diversity of commercial hybrids. Pre-breeding is the long-term conservation and utilization of genetic resources linked to an efficient cultivar development process (Carena et al., 2009b). Eleven Latin American countries and the USA created the Latin American Maize Project (LAMP) in order to evaluate maize landrace collections and identify the most useful and elite accessions for breeding programs (Pollak, 2003). Genetic enhancement efforts for their specific adaptation to target US environments and their successful utilization were followed by the Germplasm Enhancement Maize (GEM) project and its network of private and public cooperators. GEM was organized to be a public–private effort to enhance the best maize accessions identified by LAMP, by conducting a coordinated breeding effort using adapted by exotic breeding crosses to develop early-generation breeding. The NDSU maize breeding program initiated the EarlyGEM program in 1999 focused on increasing the genetic diversity of hybrids in the US North Central Region. EarlyGEM was created as a continuous
1.1 Quantitative Genetics

Basic research to formalize maize breeding procedures was stimulated by the introduction of quantitative genetic theory to breeders. Researchers at North Carolina State University (NCSU) were responsible primarily for applying quantitative genetic theory to maize breeding during the latter part of the 1940s. Quantitative genetics had been used extensively by animal breeders, but plant breeders had used only limited quantitative genetic theory for guidance in the development of breeding plans. The emphasis given to inbred line and hybrid development and the large numbers of genotypes with a relatively short-generation interval that could be evaluated led breeders to ignore the potential of quantitative genetics. Since 1950 quantitative genetics has strongly influenced the thinking in maize breeding. Population structure, population improvement based on structure, selection theory and methods, types of gene action, predicted genetic gain, and mating designs were concepts that became known to the maize breeder and served as guides in developing long-term breeding programs on genetically broad and narrow breeding populations. In addition to the principles and terminology of quantitative genetics, an important aspect (previously generally neglected) was the importance of improving the basic breeding populations, i.e., population improvement via recurrent selection during pre-breeding. The first population improvement programs usually were initiated to compare empirical results with theoretical predictions. Because of the seeming yield plateau in the 1950s, population improvement programs seemed logical for upgrading the breeding populations used as source materials for extraction of inbred lines. Estimates of genetic components of variance showed that a preponderance of total genetic variance was additive. Hence, cyclical population improvement (generally called recurrent selection) programs seemed to be an obvious method for increasing the frequency of superior lines while maintaining genetic diversity. B73 is a good example on how a cyclical half-sib recurrent selection program improved a genetically broad-based population as a source of new and elite inbred lines.

General acceptance of the principles of quantitative genetics in maize breeding was relatively slow by applied maize breeders, whose major responsibilities were for development of elite inbred lines for use in hybrids. Most applied breeders, however, are cognizant of quantitative genetics research and usually appreciate its importance for quantifying the genetic parameters they frequently have observed empirically. Reviews by Dudley and Moll (1969), Moll and Stuber (1974), and Hallauer (1992) have summarized and interpreted data obtained from quantitative genetic studies and how they are related to plant breeding. More extensive quantitative genetic studies have been conducted with maize than with most other crop species. In maize one
has greater flexibility of mating, and seed supplies usually are not limiting, which makes maize amenable for development of progenies for testing and recombination. Several progeny types and methodologies are available to even increase the amount of seed needed. However, recent efforts on QTL identification on mapping populations have focused on relatively smaller sample sizes than often produced in classical quantitative genetic studies. Thus, few efforts of integrating results from both methodologies have been reported. Epistasis and heterosis are good examples of the limitations still present with both methodologies.

The methods of maize breeding depend on the objectives of the breeder, demands of the seed industry, and the level of agricultural development of the areas in which the final product is to be used. The objectives of the breeder may be long-, intermediate-, and short term, in which case the objectives may vary at different stages of the breeding program (Fig. 1.3).

Breeding programs need flexibility to meet future contingencies that are either predictable or non-predictable, e.g., the southern leaf blight \((\text{Helminthosporium maydis})\) epidemic in the USA in 1970. In certain developing countries, an inbred line and single-cross hybrid breeding program may have low priority because of the problems associated with costs for the production and distribution of seed. Breeding programs in developing countries may have objectives of developing, evaluating, and improving populations and testing population hybrids; these objectives may have been initiated in more advanced programs several years ago and may have little interest for breeders emphasizing inbred line and hybrid development unless unique inbred lines are to be developed. All programs, however, should have long-, intermediate-, and short-term objectives to have current relevance and long-term viability for the sustainability of maize production in a determined area. The breeder should have germplasm identified for present objectives and areas (e.g., breeding and testing, not only testing in the area targeted and not elsewhere) and be working on germplasm for intermediate and future objectives.

Fig. 1.3  Objectives of breeding programs to meet present and anticipated goals
1.2 Population Improvement: What Do We Mean by Recurrent Selection?

Table 1.2 Methods of maize selection procedures that have been suggested and used for inbred line development and population improvement

<table>
<thead>
<tr>
<th>Population improvement</th>
<th>Intra-</th>
<th>Inter-</th>
<th>Inbred line development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mass (M)</td>
<td>1. Mass (M)</td>
<td>1. Pedigree</td>
<td></td>
</tr>
<tr>
<td>b. Testcross (HT)</td>
<td>3. Full-sib (FS)</td>
<td>4. Gamete selection</td>
<td></td>
</tr>
<tr>
<td>3. Full-sib (FS)</td>
<td>4. Inbred progeny (S)</td>
<td>5. Monoploids</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Androgenesis – paternal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Gametophyte factor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. Pollen culture</td>
<td></td>
</tr>
</tbody>
</table>

Several distinct methods of maize breeding techniques have evolved with the transition from the growing of open-pollinated varieties to hybrids (Fig. 1.1). The chronological phases are not distinct because of considerable overlapping, personal preferences, data and information available as maize breeding developed, and objectives of the breeding programs. Any listing of the methods of maize breeding becomes arbitrary because they may or may not be applicable for short- vs. long-term objectives, intra- vs. inter-population selection procedures, and population improvement vs. inbred line and hybrid development and/or both together. The methods are listed under two broad categories: (1) inbred line development for eventual use in hybrids and (2) improvement of populations for use either as breeding populations and/or sources of inbred lines or by farmers (Table 1.2). The stage of development of maize breeding, population genetic structures, and husbandry will determine which method(s) in either category is used. Obviously, some of the methods in each category and more than one method in one category are used in specific breeding programs. Programs that have short- and long-term objectives would include methods from both ideally integrating pre-breeding with cultivar development.

1.2 Population Improvement: What Do We Mean by Recurrent Selection?

Population improvement provides for the cyclical upgrading of open-pollinated varieties, synthetic varieties, and composites formed from a mixture of races, varieties, and inbred lines. Often, recurrent selection is claimed as a means for recycling inbred lines developed through pedigree selection of elite × elite combinations. The
term ‘recurrent selection’ was suggested by Jenkins (1940) for one method of intra-population population improvement (e.g., based on S1 lines) and later described for population improvement with the use of a tester (Hull, 1945; Hallauer et al., 1988). Recurrent selection was later redefined as a group of breeding procedures consisting of recurrent cycles of selection for outstanding genotypes with a specific purpose in a heterozygous population and the subsequent recombination of the selected portion of the population (Lonnquist, 1952). They were promoted to solve the ‘yield ceiling’ reported in the 1950s as a means to improve the grain yield of inbred lines. The greatest virtue of recurrent selection methods is that they increase the frequency of favorable alleles for traits that are quantitatively inherited while maintaining genetic variability for continued genetic improvement (Jenkins, 1940; Hull, 1945; Comstock et al., 1949; Horner, 1956).

Maize breeders in the private sector have used recurrent selection in a limited way (Weatherspoon, 1973; Good, 1990). There was an attempt to compare various recurrent selection methods with the pedigree method of developing inbred lines (Duvick, 1977; Good, 1990). However, the development of elite inbred lines depends on the improvement of germplasm sources that may include both genetically narrow- and broad-based populations. An important similarity to note is that both methods use evaluation strategies (see Fig. 1; Carena and Wanner, 2009). Both breeding methods need to be complementary and will only be successful if both are given the same importance (Hallauer, 1985). Recurrent selection studies in maize were developed to obtain improved sources of germplasm for the potential extraction of inbred lines and their possible use in hybrids. Continuous germplasm improvement is very valuable as a source of new and diverse maize hybrids that ideally have good combining ability across heterotic groups (Barata and Carena, 2006). The industry still needs inbred lines that are new, different, and unrelated to existing lines (Weatherspoon, 1973) and the sustained future success of the hybrid industry depends on increasing the genetic potential of germplasm (Coors, 1999). Use of elite and diverse germplasm should increase the rates of average grain yield improvement in the 21st century.

Improvement of populations may be either for use as source populations for new inbred lines or for use by the farmer as populations per se or in hybrid combinations. The former is for use in more advanced programs, whereas the latter is for use in certain developing areas or countries that do not have the means for production, distribution, and growing of single-cross hybrids or simply as a means for an alternative low-cost system. Hence, population improvement has always been an important objective from the most primitive to the most advanced maize breeding programs. Intra-population improvement is as old as maize breeding, whereas inter-population improvement was stimulated by the expression of heterosis.

Methods of selection for population improvement have evolved from the simplest type of mass selection for intra-population improvement to the complex procedures of reciprocal recurrent selection for inter-population improvement. Obviously, all methods have been successful for some traits at some stage of maize breeding development. Some evidence has been presented for the relative efficiency of the different methods of selection, but additional data are still needed.
to determine their long-term effects. Some of the methods are applicable only to intra-population improvement, whereas others are useful for either intra- or inter-population improvement (e.g., half-sibs and full-sibs). The choice of any one method of selection depends on the breeder, stage of the breeding program, stage of germplasm development, stage of knowledge of the populations, and objectives of the breeding program. Usually each breeding program will include more than one method of selection either for one population or for more than one. Cyclical breeding methods usually are associated by genetically broad-based populations, but similar concepts are used in pedigree selection programs (Hallauer, 1985). Major difference, but not necessarily for either genetically broad- or narrow-based populations, is that closed populations is the rule for genetically broad-based populations, whereas different germplasm sources are introduced to improve specific traits within pedigree selection programs. Fig. 1.4 shows an example of the type of germplasm sources that can be used for a maize breeding program.

### 1.2.1 Mass Selection

Mass selection is undoubtedly the oldest method used in maize breeding. The selection unit comprises the non-replicated phenotypes of individual plants. Ears of the selected plants are harvested and shelled and equal quantities of the selected seed are composited for planting the following season. Mass selection was effective in fixing some traits, but it generally was not effective for yield improvement (Sprague, 1955) because of the techniques frequently used. Factors that contributed to the seeming ineffectiveness of mass selection were poor isolation, no environmental

---

**Fig. 1.4** Germplasm pre-breeding sources for NDSU maize inbred line development (adapted from Carena et al., 2009c)
control, genotype–environment interaction, no parental control, and poor plot technique. The distinctive differences among the many different populations of maize, however, show that mass selection was effective for fixing traits that are not significantly affected by the environment (e.g., days to silking). Mass selection, therefore, was effective for traits of high heritability. Some of the disadvantages of mass selection as an effective breeding method were corrected by the suggestions of Gardner (1961). Although mass selection is the oldest method of maize breeding, it is still used for the improvement of breeding populations; its effectiveness is dependent on the plant trait selected, adequate isolation, sample size utilized, and the precision of the experimental techniques used by the breeder. This method allows two major advantages: selection at one season per cycle with no limit on population sizes. Carena et al. (2008) and Hallauer (1999) have shown results of up to $-3.8$ days of silking per year of genetic improvement in earliness with up to 22,500 F$_2$ individuals screened each cycle.

1.2.2 Half-Sib Selection

Half-sib selection utilizes, in all instances, a common tester parent. The tester parent could be either the population that is under selection (intra) or a genotype unrelated to the population under selection (inter); one that is either a narrow genetic base (inbred line) or a broad genetic base (open-pollinated, synthetic, and composite varieties). The tester can be one that either has poor or has good combining ability. Half-sib selection probably has been used more frequently than any other method of selection of population improvement.

Choice of a tester parent for use in half-sib selection depends on the objective of the breeding program and knowledge of the types of gene action operative in yield heterosis. Individual plants of the population under selection or lines derived from it are crossed with a common tester; this may be done either by hand pollination (where the tester is the female parent and males can be used for crossing and selfing for advancing a generation at the same time) or in an isolation plot where the tester is the male parent and females are detasseled. If the tester is not an inbred line, a sister line, or a hybrid, adequate sampling of the gametic array of the tester is also necessary. The half-sib progenies (all progenies that have a common parent) are evaluated in replicated multi-environment yield trials to determine the superior performing genotypes. Either remnant seed of the half-sib progenies (when the parental population is used as the tester) or selfed seed of plants used in making the crosses are recombined to form the next cycle population. Half-sib selection can be used for intra-population improvement or for reciprocal recurrent selection as suggested by Comstock et al. (1949).

1.2.2.1 Ear-to-Row

The ear-to-row selection method is, with adequate isolation, a form of half-sib selection. The method was devised and introduced by Hopkins (1899) for improvement
of the chemical composition of the maize kernels in the oldest long-term selection experiment in the crop. Ears from a population are planted ear-to-row and evaluated in a progeny row test. In contrast to mass selection, the selection unit of ear-to-row is the progeny (a half-sib family of an ear) rather than individual plants. Ear-to-row selection also requires adequate isolation to prevent contamination from other varieties. Ear-to-row selection was used to some extent by maize breeders until about 1925; it was effective in some instances and not in others depending to some extent on the trait selected and how it was affected by the environment. The ear-to-row selection method resulted in rapid improvement on non-adapted varieties in ND (Olson et al., 1927). However, no yield improvements were obtained by continuous ear-to-row selection. Two modifications of ear-to-row selection were made by Lonnquist (1964) and Compton and Comstock (1976) to make it more effective. The modified ear-to-row selection method has progeny row tests replicated in different environments, one of which is isolated from other maize fields for recombination. As a consequence, modifications increased the precision of progeny row tests and prevented contamination from other maize populations of selected progenies, thus, generating better results (Carena, 2005b).

1.2.2.2 Testcross

Selection based on testcrosses is commonly used in maize improvement, but it is a method based on half-sib family selection. Because evaluation of testcrosses is commonly used in hybrid testing and identification, the terminology may be more common than that of half-sib family selection. Expected genetic advance is the same for testcross and half-sib selections when the tester is the parental population (Sprague and Eberhart, 1977). A common tester parent is used to produce the testcross (or half-sib) progenies for evaluation and selection. Remnant seed of selected progenies is used for recombination if the tester parent is used as the male. If individual plants in the population under selection are used as males, they can be selfed and $S_1$ or $S_2$ seed used for recombination to form the succeeding cycle population.

Testcross selection usually was used for intra-population improvement only, but Russell and Eberhart (1975) proposed the use of inbred testers as a modified form of reciprocal recurrent selection. In place of using the opposing populations as testers, they suggested the use of elite inbred lines, one from each of the two populations under selection. The suggestion of currently usable elite inbred lines as testers was predicated on the evidence that the greatest proportion of the total genetic variation arises from additive genetic effects. Because additive gene action seems to be of major importance, substitution of inbred lines as testers should not hinder future progress from selection. Also, such use as testes to produce the testcrosses (or half-sibs) for two opposing populations would be simpler than the half-sib selection of reciprocal recurrent selection. Testcross selection has been used more extensively for intra-population improvement but seems to have promise for inter-population improvement as well. The use of industry testers from opposite heterotic groups has also promised as an additional modification to the system as future lines developed from improved population would have certain affinity to the testers already.
1.2.3 Full-Sib Selection

Unlike half-sib selection (only one parent in common) full-sib selection evaluates progenies having the same two parents in common. Full-sib progenies usually are developed by crossing two individuals that are either in the same population (intra) or in two separate populations (inter). Full-sib progenies are evaluated in replicated multi-environment yield tests and remnant seed is used for recombination to form the next cycle population. Full-sib selection, however, has not been used by maize breeders to the same extent as half-sib selection. A possible explanation is the amount of seed that can be produced for extensive testing. Hallauer and Eberhart (1970) suggested the use of prolific plants for producing full-sib progenies; full-sib progenies are produced reciprocally on one ear of each plant, and the other ear on each plant is selfed and used for recombination to form the next cycle population. In addition, Hallauer and Carena (2009) have proposed the use of S1 lines for increasing the amount of full-sib progeny seed for testing and reducing $G \times E$ interactions in which winter nurseries would be required.

1.2.4 Inbred Progeny Selection

The inbred or selfed progeny method of selection has not been used and tested as extensively as mass, half-sib, and full-sib selection for population improvement. Selfed progeny selection includes systems that evaluate selfed progenies themselves, uses the information obtained in multi-environment trials to determine the superior progenies, and recombines remnant seed of the selfed progenies to form the next cycle population. The generation of inbreeding of the progenies evaluated usually is in the $S_1$ or $S_2$ stages, although any level of inbreeding could be used. The advantages of selfed progeny selection are increased variability among the progenies evaluated, exposing of the deleterious recessive genes that can be eliminated, and advanced generations ready for pedigree selection. Disadvantages may be longer cycle intervals (which can be minimized by off-season nurseries), greater experimental errors in progeny evaluation, and possible linkage and inbreeding effects if advanced generation lines are recombined. Sezegen and Carena (2009) proposed the ‘intra-diallel’ recombination method for selecting and recombining $S_1$ progeny lines to achieve one season per selection cycle when selecting for cold tolerance.

1.3 Inbred Line Development

The most successful maize inbred line, B73 (Russell, 1972), was a diverse product developed in the public sector and generated billions of dollars to industry and farmers. Its germplasm source was an improved version of Iowa Stiff Stalk Synthetic or BSSS (Sprague, 1946), a genetically broad-based population that was improved by five cycles of half-sib recurrent selection. Even though the odds of developing
successful public lines from genetically broad-based improved populations seem to be low, it only requires one to make a large and significant impact (Hallauer and Carena, 2009). Public maize breeding programs provide genetic diversity in reserve (Duvick, 1981). Moreover, future genetic gains are still dependent on the deployment of useful genetic diversity carried out in the public sector (Smith, 2007) and cultivar development in areas not heavily invested in by the industry. Northern and western North Dakota environments are clear examples of the need for on-site cultivar development programs (e.g., besides North Dakota State University public program) focused on drought tolerance, cold tolerance, earliness, and fast dry down.

Because of the acceptance of the use of inbred lines in production of hybrids, breeding methods that emphasize line development are used extensively (Fig. 1.1). Other breeding methods often were neglected and emphasis was given to isolating and recycling inbred lines and testing them in hybrid combinations.

### 1.3.1 Pedigree Selection

Pedigree selection and its modifications are the most commonly used selection methods of inbred line development. Pedigree selection became common after Shull (1908, 1909) outlined the principles of inbreeding and hybridization and when industry exploited the practical value of heterosis. Progenies are developed from source populations by some form of inbreeding, with selfing as the most commonly used. Although pedigree selection often implies selection in an F2 population produced from a planned cross of two related genotypes (e.g., inbred lines in maize), pedigree selection could be practiced with progenies developed in improved open-pollinated, composite, and synthetic varieties; backcross populations; and mixtures of germplasm as well as F2 populations. Often, public maize breeding programs utilized a diverse set of improved breeding sources (Fig. 1.4). However, the most common procedure is to utilize F2 genetically narrow-based populations made from elite × elite F1 hybrids of related lines within heterotic groups.

In each instance, progenies are established that have an assigned designation (pedigree) to identify the numerical sequence (ID), generation of development (pedigree), origin (source), and remarks (Fig. 1.5). Detailed record keeping is required to properly identify the progenies selected in each generation of inbreeding across years. The selection unit usually is a combination of progeny row and individual plants within a progeny row. Seed of selected plants within selected progeny rows is planted ear-to-row the following season. Progeny rows usually are non-replicated, but because of controlled pollinations, selection is effective for traits of relatively high heritability. However, early-generation progenies can be replicated across locations within regions (Carena and Wanner, 2009) for increasing visual selection accuracy (e.g., screening for cold tolerance traits, duplicating progenies in disease or testcross nurseries). Different planting densities can be used for exposing progenies to different environments as well as producing larger seed amounts per progeny. For traits such as yield of progenies themselves and in crosses,
<table>
<thead>
<tr>
<th>ID</th>
<th>PEDIGREE</th>
<th>SOURCE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>08WNNDNZ1650</td>
<td>BS21(R-T)C8-89-2-1-1</td>
<td>083602-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1651</td>
<td>BS21(R-T)C8-89-3-1-1</td>
<td>083606-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1652</td>
<td>BS21(R-T)C8-89-3-2-1</td>
<td>083607-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1653</td>
<td>BS21(R-T)C8-89-3-2-2</td>
<td>083607-2</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1654</td>
<td>BS21(R-T)C8-89-3-2-4</td>
<td>083725-1</td>
<td>Female</td>
</tr>
</tbody>
</table>

60 NDSU Lines × TR3030

<table>
<thead>
<tr>
<th>ID</th>
<th>PEDIGREE</th>
<th>SOURCE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>08WNNDNZ1655</td>
<td>NDL(FS)C1-208-3-1-2-1</td>
<td>083452-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1656</td>
<td>(ND08-144) [(ND278 × ND290)];-38-2-1-3-1</td>
<td>083455-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1657</td>
<td>BS21(R-T)C8-89-2-1-1-1</td>
<td>083595-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1658</td>
<td>BS21(R-T)C8-89-1-1-1</td>
<td>083610-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1659</td>
<td>BS21(R-T)C8-89-2-1-1-1</td>
<td>083614-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1660</td>
<td>BS21(R-T)C8-7-1-2-1-1</td>
<td>083631-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1661</td>
<td>BS21(R-T)C8-99-2-1-1-1</td>
<td>083636-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1662</td>
<td>BS21(R-T)C8-143-1-1-2-1</td>
<td>083639-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1663</td>
<td>BS21(R-T)C8-143-1-2-1-1</td>
<td>083640-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1664</td>
<td>BS21(R-T)C8-135-1-1-1</td>
<td>083641-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1665</td>
<td>BS21(R-T)C8-135-1-2-1-1</td>
<td>083645-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1666</td>
<td>BS21(R-T)C8-33-1-1-1</td>
<td>083648-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1667</td>
<td>BS21(R-T)C8-33-2-1-1-1</td>
<td>083651-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1668</td>
<td>BS21(R-T)C8-41-1-1-1</td>
<td>083653-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1669</td>
<td>BS21(R-T)C8-41-1-2-1</td>
<td>083654-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1670</td>
<td>BS21(R-T)C8-41-2-1-1</td>
<td>083656-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1671</td>
<td>BS21(R-T)C8-41-3-1-1</td>
<td>083657-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1672</td>
<td>BS21(R-T)C8-139-1-1-1</td>
<td>083658-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1673</td>
<td>BS21(R-T)C8-139-2-1-1</td>
<td>083660-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1674</td>
<td>BS21(R-T)C8-113-3-1-1</td>
<td>083662-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1675</td>
<td>[ND2000 × ND290];-169-2-1-1-1</td>
<td>083871-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1676</td>
<td>[ND2000 × ND290];-180-1-1-1-1</td>
<td>083872-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1677</td>
<td>[ND2000 × ND290];-26-1-1-1-1</td>
<td>083873-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1678</td>
<td>[ND2000 × ND290];-62-1-1-1-1</td>
<td>083874-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1679</td>
<td>[ND291 X ND2000];-141-1-1-1-1</td>
<td>083885-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1680</td>
<td>[ND291 X ND2000];-50-2-1-1-1</td>
<td>083886-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1681</td>
<td>NDBS1011(FR-M-S)C4-49-2-1-1-1</td>
<td>083889-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1682</td>
<td>NDBS11(FR-M-S)C3-126-1-1-1</td>
<td>083890-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1683</td>
<td>NDBSK(HI-M-S)C3-45-1-1-1-1</td>
<td>083892-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1684</td>
<td>NDBSK(HI-M-S)C3-47-1-1-1-1</td>
<td>083893-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1685</td>
<td>[((Mo17 X ND2000) X ND2000);-1]-125-1-1-1-1</td>
<td>083895-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1686</td>
<td>[((Mo17 X ND278) X ND278);-2]-107-1-1-1-1</td>
<td>083896-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1687</td>
<td>[((Mo17 X ND278) X ND278);-2]-2-1-1-1-1</td>
<td>083897-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1688</td>
<td>[((Mo17 X ND278) X ND278);-2]-86-1-1-1-1</td>
<td>083898-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1689</td>
<td>[((Gem2 X ND2000) X ND2000);-1]-102-1-1-1-1</td>
<td>083899-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1690</td>
<td>[((Gem2 X ND2000) X ND2000);-1]-8-2-1-1-1</td>
<td>083900-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1691</td>
<td>[((Gem2 X ND2000) X ND2000);-1]-94-1-1-1-1</td>
<td>083901-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1692</td>
<td>[((Gem2 X ND2000) X ND2000);-1]-15-1-1-1-1</td>
<td>083902-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1693</td>
<td>[((Gem2 X ND2000) X ND2000);-1]-79-1-1-1-1</td>
<td>083903-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1694</td>
<td>[ND291 X B100];-42-1-1-1-1</td>
<td>083904-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1695</td>
<td>[ND291 X B100];-59-1-1-1-1</td>
<td>083905-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1696</td>
<td>[ND291 X B100];-6-1-1-1-1</td>
<td>083906-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1697</td>
<td>[ND291 X B100];-62-1-1-1-1</td>
<td>083907-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1698</td>
<td>[ND291 X B100];-67-1-1-1-1</td>
<td>083908-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1699</td>
<td>[A619E X ND291];-27-1-1-1-1</td>
<td>083909-1</td>
<td>Female</td>
</tr>
<tr>
<td>08WNNDNZ1700</td>
<td>[A619E X ND291];-5-1-1-1-1</td>
<td>083910-1</td>
<td>Female</td>
</tr>
</tbody>
</table>

**Fig. 1.5** An example of pedigree records for S4 and S5 non-stiff stalk lines that survived early-generation testing and are now being crossed with industry tester TR3030 (Thurstons Genetics early-maturing SS tester)
multi-environment replicated yield tests become necessary at some stage of the pedigree selection; in most instances testing is done in the early generations of inbreeding complemented to late-generation hybrid testing of selected progenies approaching homozygosity. Pedigree selection is extensively used in recycling of lines that have known strengths and weaknesses for specific traits. Pedigrees developed by recycling of lines can become quite complex (Fig. 1.6 and Table 1.3), but parentage control permits the development of lines to meet specific requirements. Techniques of pedigree selection are described by Harrington (1952), Allard (1960), Hallauer (1987) and used extensively in maize breeding programs (Bauman, 1981; Bernardo, 2006; Carena and Wanner, 2009; Hallauer and Carena, 2009; Carena et al., 2009a–c).

### 1.3.2 Backcross Selection

The backcross method is a special case of pedigree selection which it is seldom used in breeding but often used for transgenic trait integration. Planned crosses are made, as with pedigree selection, but usually it is desired to incorporate a specific gene(s) into an otherwise desirable genotype either with molecular markers or without them.

![Fig. 1.6](image-url) An example of pedigree selection that illustrates the principle of recycling in the development of inbred lines. I205 and Idt (I205-5-4) are sister selections from Iodent developed by L. C. Burnett (Wallaces, 1923); B164 (or Ind 461-3) is from Troyer Reid; LLE is from a Long Ear line developed by E.W. Lindstrom (Troyer, personal communication); and M49 is the Minnesota culture of 1949 [information courtesy of Mr. Raymond Baker, Pioneer Hi-Bred International, Inc., Johnston, Iowa, with additions by Troyer (1999)]
Table 1.3 Recycling of Iowa State and North Dakota State maize inbreds

<table>
<thead>
<tr>
<th>Source</th>
<th>Original</th>
<th>Derived (selected examples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS B14</td>
<td>A632, A636, CM105, B64, B68, FR140, LH190, HC60, LH224, LH225, LH301, MBS1130, TR1017, TR2040, TR3026, TR3030, TR4033</td>
<td></td>
</tr>
<tr>
<td>B37</td>
<td>B76, H84, H93, LH195, TR3030, TR3273, SGI890, SGI912, LH132, LH225, HC34, MBS2681</td>
<td></td>
</tr>
<tr>
<td>B73</td>
<td>A681, N196, N201, NC292, H123, FS1064, LH195, LH198, TR3026, TR3127, SGI890, SGI952</td>
<td></td>
</tr>
<tr>
<td>Minnesota #13</td>
<td>ND203, ND230</td>
<td>A624, A640, A641, A664, A665, LH201, LH147, SD106, PA3401, PA3402, PA3422, PA3432</td>
</tr>
<tr>
<td>W755 × W771</td>
<td>ND246</td>
<td>LH160, LH161, LH162, LH256</td>
</tr>
<tr>
<td>B14 × WD</td>
<td>ND302</td>
<td>A635, A636, FR19</td>
</tr>
<tr>
<td>NDSCD(M)C8</td>
<td>ND2000</td>
<td>NDSU EarlyGEM experimentals</td>
</tr>
</tbody>
</table>

In maize breeding the objective of backcross selection often is to transfer a specific trait (e.g., Ht, o2, fl2, lg2, Lg3, and/or lately Bt, RR with marker-assisted backcrossing) from a donor parent (non-recurrent) to an elite inbred line (recurrent parent); the breeder wants to recover the genotype of the elite inbred line in successive backcrosses and incorporate the desired trait from the donor parent. As for pedigree selection, detailed records are needed to identify recurrent parent, donor parent, generation of backcrossing, and specific trait under selection (Fig. 1.5). Because the trait transferred by backcrossing usually can be classified in discrete classes, the selection unit is a combination of individual plants within non-replicated progeny rows. To transfer a one-gene trait from one parent to another, backcross selection has an advantage over pedigree selection because 50% of the plants for backcross selection would be homozygous for the trait compared to only 25% of the plants for pedigree selection in the first selfing generation. Backcrossing also may be used for traits that are not controlled by one or two genes, e.g., incorporation of exotic germplasm into adapted populations. The NDSU maize breeding program has developed BC1 populations derived from elite early × late crosses to adapt GEM breeding crosses in the northern USA and to increase the genetic diversity of industry early-maturing hybrids through the NDSU EarlyGEM project (Carena et al., 2009b). Early × late crosses have also been used successfully to develop earlier versions of successful later-maturing lines (see table 8.3 of Hallauer et al., 1988). In these instances larger plant numbers are needed to sample the array of genotypes in the recurrent and non-recurrent parental populations. Backcrossing is a convenient method for either the transfer or the improvement of a specific trait into an elite inbred line. Hallauer et al. (1988) in their Table 8.3 presented a summary of successful approaches toward developing early-maturing lines from early by late crosses. Richey (1927) developed
the concept of convergent improvement, in which single-cross hybrids are backcrossed to both parental lines in order to accumulate all dominant favorable genes. Convergent improvement, as for backcrossing, has been successful for introducing simply inherited traits into established lines; but for complex traits such as yield the method is less efficient. The techniques and principles of backcrossing are discussed in greater detail by Hayes et al. (1955) and Allard (1960). Convergent improvement has been less important than backcrossing in maize breeding.

1.3.3 Single-Seed Descent

In spite of its advantages the single-seed descent method is not used extensively in maize breeding. Jones and Singleton (1934) suggested a modification called the single-hill method but the method has not been used to any great extent in maize either. The single-seed descent method is not as commonly used in maize breeding as in self-pollinated species. This method is used when a population of genotypes is sampled and at least one seed is saved from each genotype in successive generations of inbreeding until approximate homozygosity is reached. The same ear can be used (e.g., cut) for different methods of inbreeding. The clear advantage of this method is the smaller nursery space needed compared to the pedigree method of inbreeding while trying to reach homozygosity faster with minimum resources. Except for natural selection a minimum of selection pressure is applied. At some level of homozygosity, seed supplies of the progenies obtained by single-seed descent are increased for evaluation in replicated tests. For obligate self-pollinated species, self-pollinations are made naturally.

The pedigree and backcross methods use the principles of the single-seed descent method because selected plants in selected progenies are propagated to the next generation. The distinction is that artificial selection is maximized in the pedigree and backcross methods and minimized in single-seed descent. Natural selection, of course, is operative in all methods. The single-seed descent method has been used in maize for quantitative genetic studies estimating parameters in reference to a specific population (Hallauer and Sears, 1973; Silva and Hallauer, 1975). Because controlled pollinations are required to affect inbreeding, the single-seed descent method is not as amenable for maize as for self-pollinated species.

1.3.4 Gamete Selection

Gamete selection was suggested by Stadler (1944) and, as single-seed descent, has had limited use in maize breeding when compared to the pedigree and backcross selection methods. Stadler theorized that if the frequency of superior zygotes was $p^2$, the frequency of superior gametes was $p$, i.e., if $p^2$ is 0.25, $p$ is 0.5. Because the theoretical frequencies of superior gametes were greater than for superior zygotes, gametic sampling should be more efficient than zygotic sampling. Briefly, gamete
selection includes crossing a bulk sample of pollen from a source population with an elite inbred line in season 1, self-fertilizing the F₁ plants (each F₁ plant differing only by the gametic complement contributed by the source population) and crossing the F₁ plants and an elite inbred line to a common tester in season 2, and growing the testcrosses in replicated yield trials in season 3. Therefore, any testcross that exceeded the elite line × tester combination presumably received a superior gamete from the source population. Although the testcrosses identify the superior gametes, the major disadvantage is that the superior gamete cannot be fixed in homozygous inbred lines. Because individual plants by elite line F₁ plants were self-fertilized, only 25% are homozygous for individual loci; this would increase to 37.5% over all S₁ progenies, but not all progenies would be retained from the testcross data. Selection units are individual plants from the source population but would include progeny row and individual plants within the selfed progenies of the F₁ plants. Some positive data have been reported from use of gamete selection, but the method has not been used generally. The method seems to have merit and with some modifications could be useful for identifying superior genotypes in a source population.

Zygote selection was suggested (Hallauer, 1970) as one modification of gamete selection. Plants in a source population are selfed and crossed with an elite line used as a tester. The gametic array of the plant used in test-crosses would be similar to that of the S₁ progenies. Although recombination would prevent the recovery of identical superior gametes expressed in the testcrosses, the opportunities would be greater than by gamete selection. In all instances the primary objectives of these methods are either to improve an elite inbred line while development of derived ones and/or to select a companion line for use in hybrids.

Certainly a good example of gamete selection is the gametophytic system invention which limits foreign pollen from contaminating seed produced through pollen–pistil interactions (e.g., US patent 6875905 by Thomas C. Hoegemeyer), a major concern lately with transgenic products.

1.3.5 Monoploids

The fast development of homozygous plants has always been of interest to maize breeders, especially for new breeding programs. Several interesting methods for rapid development of inbred lines have been suggested. In most instances methods have neither been adequately evaluated to determine its relative merits nor been assessed to determine their efficiency in developing elite inbred lines. Development of inbred lines, however, is generally not a serious problem in maize breeding especially in established breeding programs that provide lines at different generations of inbreeding yearly. Determining the relative merit of the lines in hybrids is the more difficult problem. Also, tremendous selection pressure can be applied during inbreeding for the phenotypic expression of plant and ear traits, disease and insect resistance, and yield in crosses by use of early-generation testing. Techniques
for rapid development of inbred lines yield a group of lines that are a random sample for most traits, several which could be discarded earlier. After the inbred lines are available, they have to be grown and evaluated for the traits necessary for use as parental seed stocks in hybrids increasing the number of hybrid combinations needed from unselected inbred lines.

1.3.5.1 Homozygous Diploids

Production of homozygous diploids of maize from monoploids (e.g., haploids) was outlined by Chase (1952a, b) as a new technique for inbred line development (e.g., doubled haploids). Use of the method depends on the occurrence of monoploids and maternal monoploids occurred spontaneously at a high enough rate \( \left( \frac{1}{1000} \right) \) to yield sufficient numbers of monoploid plants. Successful production of homozygous diploids from monoploids depends on the production and recognition of monoploids and on the deriving homozygous, diploid progeny derived from the isolated monoploids. Spontaneously occurring monoploids were identified by Chase (1949) in the seedling stage by the use of a genetic marker system, the dominant allele for purple plumule (Pu). Many of the monoploids survived to maturity to produce normal gametophytes. Successful self-pollination of the mature monoploids yielded homozygous, diploid lines at a frequency of one homozygous diploid line for every 10 isolated monoploids. At this time, opportunities for selection have occurred only for the self-pollinated monoploid plants.

Relative merits of the homozygous diploid method for practical breeding were tested by Thompson (1954) by comparing the combining ability of yield of selected and unselected lines developed by the usual inbreeding system of self-pollination with lines developed from monoploids; all lines originated from the same source population BSSS. The lines derived from monoploids were a random sample with respect to yield in topcrosses because the differences between means were non-significant and the frequency distributions were similar. Even though few lines were developed this way (e.g., B67 and B69) the method has not been used extensively by maize breeders. Theoretically, the homozygous diploid method has the same advantage as Stadler’s (1944) gamete selection method in which the frequency of superior gametes is greater than the frequency of superior zygotes. The advantage of this method, however, is that the superior gametes of monoploids can be fixed in the homozygous condition in just one generation of self-pollination. The difficulty is in identifying the superior gametes, which can be done only in observations repeated over environments.

Interest in the haploid method of maize breeding has wavered over time (e.g., Chase, 1949; Kermicle, 1969) because the main concern has been the frequencies in the occurrence of haploids (\( n = 10 \)) and being able to double the chromosome number of the haploids to produce fertile diploids (\( n = 20 \)). There has, however, been a significant resurgence in the interest and use of haploid breeding methods in maize (Seitz, 2005). The main attraction of the potential of haploid breeding methods (being the same as in the past) is the immediacy of acquiring homozygous
inbred lines (Seitz, 2005; Longin et al., 2006) confirming that haploid regeneration is an inherited trait (Spitko et al., 2006). The usual method is crossing normal maize with special stocks, designated as inducers. The pollen of inducers causes the normal types to produce seed containing haploid embryos that are chemically treated to produce fertile plants with the $2n$ complement of identical chromosomes. This method has been used since suggested originally by Chase (1949, 1952a, b), but the technology, facilities, and genetic materials (such as better inducer stocks) have improved so that a greater frequency of doubled haploids can be obtained. Developments within molecular genetics can also permit screening for identified loci at the haploid stage before continuing to the diploid stage. Although the materials and techniques have certainly been enhanced in recent years, the production of a significant number of fertile diploids to screen for performance as inbred lines and in hybrids still requires extensive human and physical resources. The time frame for obtaining homozygous inbred lines is reduced by the use of haploid breeding methods, but the final determination of their value will be similar to other breeding methods. It seems that haploid breeding methods have, and are, more important presently than they were previously. A certain advantage is the simplicity compared to an early- and late-generation testing program. Certainly, doubled haploids can be a complementary breeding strategy for germplasm improvement and inbred line development in specific populations of lines already tested for combining ability. Moreover, the frequency of haploid embryos can be significantly increased by proteins involved in growth stimulation as demonstrated by certain industry patents.

1.3.5.2 Androgenesis

Goodsell (1961) described a technique whereby paternal monoploids could be identified and selfed in one generation to produce homozygous, diploid lines. The breeding procedure is similar to the one used by Chase (1952a, b) for maternal monoploids and requires a genetic marker system to identify the monoploids. Goodsell used a genetic marker that caused purple coloration in the seedling root tips, which was similar to Chase’s system. Non-purple, fertile males were crossed with purple, male-sterile females. The F$_1$ seedlings of the crosses were examined for seedlings failing to exhibit the purple coloration of the root tips. From 400,000 progenies of four crosses Goodsell found 4 seedlings having white roots, which is lower than the frequency (1/1000) reported for maternal monoploids. Pollinations of the monoploids resulted in plants similar to the male pollinator and male sterile, i.e., the male gamete failed to unite with the female gamete but acquired the cytoplasm of the female nucleus. This method has the same advantages and disadvantages of maternal monoploids and so far has not been used to any extent. The method also seems promising for quick conversion of lines to male-sterile cytoplasms, provided one can increase the frequency of occurrence of paternal monoploids.

1.3.5.3 Indeterminate Gametophyte

Kermicle (1969) reported on a spontaneous mutation, designated indeterminate gametophyte (ig) that produced unusual effects in the embryo sacs of the plants
carrying the mutant gene. Because of the influence of the ig gene on female gametophyte development, some of the effects associated with ig are as follows: homozygous mutant plants (igig) are generally male sterile; about 50% of the kernels produced on igig plants and 25% of those on Igig plants either abort or are defective; about 6% of the seeds with normal endosperm that received ig from either Igig or igig females are polyembryonic; and when igig plants are crossed as females about 3% of the progeny are monoploids, which occurred as maternal and paternal monoploids in a ratio of about 1:2. Hence, the action of the ig gene reflects the loss of normal functions in the female gametophyte development.

It seems that production of monoploids by the ig gene may have useful advantages in maize breeding programs for (1) rapid isolation of new inbred lines by doubling of the paternal monoploids and (2) rapid conversion of existing elite lines with normal cytoplasms to versions with male-sterile cytoplasms. These two advantages are the same as those given by Goodsell (1961) except the occurrence of monoploids is at a higher frequency, which was emphasized by Goodsell if the method of androgenesis was to be useful in breeding programs. Kermicle (1969) found that the anthocyanin pigmentation characteristics of the R-Navajo (R\text{nj}) gene and its acyanic allele r\text{g} were useful genetic markers for the detection of monoploids. Reciprocal crosses of r\text{g}r\text{g} and R\text{nj}R\text{nj} stocks were used to detect monoploids of maternal and paternal origins, respectively. So far as known, the indeterminate gametophyte is not used extensively in maize breeding, but the higher frequency of monoploids would make it more attractive than any of the previously described systems.

1.3.5.4 Pollen Culture

Recent developments of tissue and cell culture techniques offer tremendous opportunities for the visionary plant breeder. If one assumes that a maize plant produces 1 – 10 million pollen grains, large numbers of haploid plants could be developed from cultured microspores. Therefore, it would be desirable to utilize techniques to culture microspores of maize to differentiated haploid plants that are doubled to homozygosity and fertile. Such techniques would provide for thousands of monoploid plants that could be doubled to homozygous, diploid inbreds. To be useful, some monitoring tests could be developed that would determine the relative merit of the doubled monoploids for disease and insect resistance, grain quality, and yield potential both as homozygous inbred lines and in hybrids. The possibilities of pollen culture could add to breeders’ choices for increasing the frequency of monoploids if practical methods for pollen culture are utilized.

1.4 Conclusions

Breeding methods (e.g., Table 1.2) for both inbred line development and hybrid identification have a place for success in maize breeding programs as long as source germplasm is well chosen. Pedigree selection has been used extensively in the past, is being used extensively at the present, and will continue to be used in the future. On the other hand, other methods, such as mass selection and backcrossing, will only be
used in more specific instances. The choice of breeding methods used in a specific program will be dictated by the circumstances at a given site. For instance, in the USA most conditions dictate that hybrids will be used by the farmer; consequently, breeding methods that enhance the development of superior inbred lines for use in the production of hybrids are of paramount importance. However, the private (Table 1.4) and the public (Table 1.5) sectors should continue to work together for the identification of superior hybrids while increasing the genetic diversity on farms through the utilization and improvement of both genetically narrow- and broad-based improved germplasm sources through not only short-term but also long-term programs.

Table 1.4  A hybrid maize commercial breeding program (adapted from Bernardo, 2002)

<table>
<thead>
<tr>
<th>Season</th>
<th>Activity</th>
</tr>
</thead>
</table>
| Summer 1 | (1) Grow 45 F2 or BC1 segregating populations from elite × elite hybrids  
(2) Select S0 plants in each population based on plant type  
(3) Self and testcross 50 selected S0 plants in each population to a tester |
| Summer 2 | (1) Discard 10% of segregating populations based on additional data performance on parents in summer 1  
(2) Grow 2000 S0 testcrosses (average of 50 per population) in unreplicated trials at 3 locations  
(3) Grow S1 progenies (ear-to-row from S0 plants selfed in summer 1). Self three plants per S1 family row to obtain three S2 subfamilies. Discard S1 progenies that visually appear inferior.  
(4) Select the best 200 S1 families (now S2 seed) based on their S0 testcross performance |
| Winter 1 | (1) Cross the best 600 S2 families to two testers  
(2) Self S2 families to obtain S3 families |
| Summer 3 | (1) Evaluate 1200 S2 testcrosses in unreplicated trials at 6–10 locations  
(2) Select 5–10 S3 families based on S2 testcross performance |
| Winter 2 | Self the selected S3 families to obtain S4 seeds of new inbreds |
| Summer 4 | (1) Cross each new inbred with 5–10 elite inbreds  
(2) Self the S4 families to obtain S5 seeds of new inbreds |
| Summer 5 | (1) Yield trials of experimental hybrids at 15–50 locations  
(2) Self the S5 families to obtain S6 seeds of new inbreds |
| Summer 6 | (1) Yield trials of advanced hybrids at 20–75 locations  
(2) On-farm strip tests (i.e., 150–300 m² plots) at 30–500 locations |
| Summer 7 | On-farm strip tests of precommercial hybrids at 50–1500 locations |
| Fall    | Release 0–2 hybrids for sell to farmers |

Table 1.5  A hybrid maize public breeding program (adapted from Carena et al., 2009c and North Dakota State University)

<table>
<thead>
<tr>
<th>Season</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter 1,2</td>
<td>Elite × elite crosses and selfing of F1 hybrids and/or F1 hybrid selfing and F2 selfing and testcrossing are conducted in the same winter due to the early-maturing nature of this program allowing three generations per year. If F2 testcrossing is in the winter we proceed as in Table 1.4. Otherwise, the program continues as follows</td>
</tr>
</tbody>
</table>
### Table 1.5 (continued)

<table>
<thead>
<tr>
<th>Season</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 1</td>
<td>(1) Grow 25–30 $F_2$ or $BC_1$ segregating populations from genetically broad- and narrow-based populations (see Fig. 1.4) under high-density stress (if during summer). Some populations represent early × late crosses to develop early versions of elite lines. Pre-breeding from long-term genetic improvement programs provide additional elite and genetically broad-based segregating populations and progenies. Examples are stratified mass selection (Carena et al., 2008), intra- and inter-population recurrent selection programs (Carena and Hallauer, 2001; Carena and Wanner, 2003; Carena et al., 2003; Hyrkas and Carena, 2005; Melani and Carena, 2005), and BC EarlyGEM programs (Carena et al., 2009b). At this moment, improved NDSU populations are released (Carena et al., 2008). (2) Select earliest and desirable $S_0$ plants in each population and testcross 50–400 $S_0$ plants per pedigree with two testers (one from each heterotic group)</td>
</tr>
<tr>
<td>Winter 3</td>
<td>(1) $S_1$-derived progenies and elite hybrids are grown and screened under drought-controlled conditions (2) Selected $S_1$ progenies are selfed (3–5 best plants). A maximum of three ears are kept per family (3) Specific drought-tolerant hybrids are tested and produced</td>
</tr>
<tr>
<td>Summer 2</td>
<td>(1) Few segregating populations are discarded based on parental data (2) Grow approx. 1,000 $S_0$ testcrosses in 1–2 rep. incomplete block lattice yield trials per pedigree at 4–6 locations (stage I) (3) Grow $S_2$ progeny rows across 2–3 locations (e.g., breeding, disease, and testcross nurseries if needed) and collect emergence and seedling vigor data across locations. Discard 10–30% (depending on season). A rank-summation index is used. Self 3–5 plants per $S_2$ family row and often keep one ear per family</td>
</tr>
<tr>
<td>Winter 4</td>
<td>(1) Cross the best 200–400 $S_2$ families to two industry testers from each heterotic group (a total of four). If season is challenging, cold storage seed could be used to speed up shipping seed (2) Re-screen lines and hybrids for drought</td>
</tr>
<tr>
<td>Summer 3</td>
<td>(1) Evaluate 200–400 $S_2$ testcrosses in replicated incomplete block yield trials at 4–10 western and northern locations (stage II) (2) Grow $S_3$ progenies in Fargo breeding nursery. Harvest three ears per line</td>
</tr>
<tr>
<td>Winter 5</td>
<td>(1) Cross the 20–40 top $S_4$ families to four industry testers from each heterotic group (2) Grow selected lines (three rows per inbred) and self-pollinate. Start checking for uniformity (3) Screen for drought tolerance</td>
</tr>
<tr>
<td>Summer 4</td>
<td>(1) Yield trials of experimental NDSU hybrids at 16–20 locations (stage III). Target ND regions and maturities (e.g., western ND for drought tolerance, northern ND for cold tolerance and yield/moisture, earliness, fast dry down, test weight, grain quality, lodging, southern ND, and central ND). Heritability indices and benchmarks are used as done for pre-breeding (e.g., used in both recurrent selection trials and early/late-generation hybrid trials)</td>
</tr>
</tbody>
</table>
Table 1.5 (continued)

<table>
<thead>
<tr>
<th>Season</th>
<th>Activity</th>
</tr>
</thead>
</table>
| Winter 6 | (1) Advanced selected families by self-pollination, check for uniformity  
(2) Use bulk seed for crossing new inbreds (5–10 S6 lines) with 12–16 industry testers (at least six from each heterotic group) depending on the availability of early-maturing testers (often limited)  
(3) New inbreds and hybrids are re-tested under drought stress |
| Summer 4 | (1) Yield trials of NDSU hybrids at >20 locations (stage IV)  
(2) Grow new inbreds in the ‘potential releases’ section of breeding nursery and re-check uniformity |
| Winter 7 | (1) Top lines are crossed with more industry testers per heterotic group  
(2) Lines are advanced and tested per and in hybrid combinations for drought stress |
| Summer 5 | (1) Yield trials of advanced hybrids at >30 locations in cooperation with industry (stage V)  
(2) Re-grow new inbreds in the ‘inbred increases’ section of breeding nursery for seed distribution and re-check uniformity  
(3) Increase seed for potential line characterization and additional hybrid combinations through Foundation Seed Companies |
| Fall | (1) Release NDSU inbreds to private and public institutions as early-maturing hybrid parents and/or breeding sources of new early-maturing lines. Once the official release is done registration at the Plant Introduction Station (Ames, IA) and National Plant Germplasm System (Fort Collins, CO) follows  
(2) Exclusive agreements with Foundation Seed Companies and others with NDSU as originator are an alternative toward generating royalties  
(3) Material Transfer Agreements (MTAs), Inbred Research Agreements, and Commercialization Agreements through NDSU Research Foundation are in place for each seed request (lots of 50 k for inbreds and lots of 200 k for improved populations) |

References


Leaming, J. S. 1883. *Corn and Its Culture*. J. Steam Printing, Wilmington, OH.


References

Chapter 2
Means and Variances

2.1 Genetically Narrow- vs. Broad-Based Reference Populations

Choice of germplasm as source of elite inbred lines is the most important decision the breeder takes. No tool or breeding methodology will be successful if a poor choice is made on source populations.

A population of maize can be characterized by the following properties: diploid (2n = 20), panmictic (random mating with more than 95% of cross-pollination), monoecious (both sexes in the same individual but in different inflorescences), a tendency for protandry, and general assumptions for no maternal effects, linkage equilibrium, normal fertilization (non-competing gametes), normal meiosis, and normal segregation.

Both means and genetic variances are important factors to consider when choosing populations to be used as sources of inbred lines and hybrids. Choosing breeding populations with a high mean performance is straightforward. However, the study of genetic variation of plant populations includes different approaches for different types of populations. The reference population of genotypes may result from genetically narrow-based populations derived from a cross between two homozygous inbred lines or from genetically broad-based populations derived from improved and/or unimproved populations. Broad-based populations can be a result of crosses among a set of homozygous inbred lines (synthetic varieties), an open-pollinated variety, or a mixture of varieties and races (composites). General theories, however, make no distinction about the origin of the population unless it does not fill some of the basic requirements.

Populations derived from crosses of two elite pure lines are commonly used in plant breeding. Consequently, we can determine the genetic composition of different generations derived from crossing two pure lines, including backcross populations. The introduction to the estimation of genetic variances in these generations has the advantages that, assuming two alleles per locus, expected gene frequencies (p and q) are known and have the same value (p = q = 0.5) for segregating loci, which makes their derivations easy to interpret unlike genetically broad-based populations. The estimation of genetic variation within genetically broad-based populations in which the allele frequencies are not known is based on mating designs to develop progenies for evaluation. These progenies are based on the genetic composition for covariance.
of relatives (see Chapter 3). Analyses of variance of the progenies derived from mating designs are used to evaluate additive and dominance genetic effects, average level of dominance, epistasis, and relative heritability as well as expected genetic gain. Public breeding programs allow growing progenies for not only estimating genetic variances but also for selection without relying on just the coefficient of coancestry. Estimating genetic variances is useful for designing breeding programs, predicting response to selection, constructing selection indices, predicting hybrid performance, and allocating breeding resources more efficiently (Bernardo, 2002).

The concepts of population means and variances in current quantitative genetics theory are based on gene effects and frequencies or, in other words, on the genetic structure of the population under study. The population structure, however, depends on several other factors such as ploidy level, linkage, mating system, and a number of environmental and genetic factors. Therefore, either some of these factors must be known or restrictions must be imposed about their effects to be able to establish a theoretical model for study.

Estimated parameters refer to a specific population from which the experimental material is a sample for a specific set of environmental conditions (Cockerham, 1963). Thus one must specify the reference population for both genotypes and environments because inferences cannot generally be translated from one population to another especially after selection. In genome-wide selection, for instance, molecular markers need to be ‘re-trained’ (Hammond, personal communication) after each time selection is conducted even within populations (e.g., across recurrent selection cycles). More detailed descriptions of the population means and variances were given by Kempthorne (1957) and Falconer (1960).

## 2.2 Hardy–Weinberg Equilibrium

Assume the reference population is in Hardy–Weinberg equilibrium. In 1908 Hardy and Weinberg independently demonstrated that in a large random mating population both gene frequencies and genotypic frequencies remain constant from generation to generation in the absence of mutation, migration, and selection. Such a population is said to be in Hardy–Weinberg equilibrium and remains so unless any disturbing force changes its gene or genotypic frequency.

This concept can be translated to a single locus as any population will attain its equilibrium after one generation of random mating. The Hardy–Weinberg equilibrium law can be demonstrated by taking one locus with two alleles (A₁ and A₂) in a diploid organism such as maize. Let us consider a population whose genotypic frequencies are as follows:

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>A₁A₁</th>
<th>A₁A₂</th>
<th>A₂A₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of individuals</td>
<td>n₁</td>
<td>n₂</td>
<td>n₃</td>
</tr>
<tr>
<td>Frequency</td>
<td>P = n₁/N</td>
<td>Q = n₂/N</td>
<td>R = n₃/N</td>
</tr>
</tbody>
</table>

\[ n₁ + n₂ + n₃ = N \]

\[ P + Q + R = 1 \]
The total number of genes relative to locus A in this population is \(2N\), i.e., two genes in each diploid individual. Thus the numbers of \(A_1\) and \(A_2\) genes are \(2n_1 + n_2\) and \(2n_3 + n_2\), respectively, and their frequencies are

\[
p(A_1) = \frac{2n_1 + n_2}{2N} = \frac{n_1 + (\frac{1}{2})n_2}{N} = P + \frac{1}{2}Q
\]

\[
q(A_2) = \frac{2n_3 + n_2}{2N} = \frac{n_3 + (\frac{1}{2})n_2}{N} = R + \frac{1}{2}Q
\]

Because gametes unite at random in a population under random mating, the genotypic array and its frequency in the next generation will be

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Male gametes</th>
<th>Frequencies</th>
<th>Male gametes</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1</td>
<td>A_2</td>
<td>p</td>
<td>p^2</td>
<td>pq</td>
</tr>
<tr>
<td>Female gametes</td>
<td>A_1</td>
<td>A_1 A_1</td>
<td>Female gametes</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>A_1 A_2</td>
<td></td>
<td>p^2</td>
<td>pq</td>
</tr>
<tr>
<td></td>
<td>A_2 A_2</td>
<td>q</td>
<td>pq</td>
<td>q^2</td>
</tr>
</tbody>
</table>

So the genotypic frequencies are \(p^2(A_1 A_1) : 2pq(A_1 A_2) : q^2(A_2 A_2)\), and this population is said to be in Hardy–Weinberg equilibrium since genotypic frequencies are expected to be unchanged in the next generation. Figure 2.1 shows the variation of genotypic frequencies for gene frequencies in the range from 0 to 1.

The Hardy–Weinberg law can also be extended to multiple alleles. In general, if \(p_i\) is the frequency of the \(i\)th allele at a given locus, the genotypic frequency array is given by

\[
\sum_i p_i^2 \quad \text{for homozygotes } (A_i A_i)
\]

\[
2 \sum_{i < i'} p_i p_{i'} \quad \text{for heterozygotes } (A_i A_{i'})
\]

Fig. 2.1 Distributions of genotypic frequencies for gene frequencies ranging from 0 to 1.0 for one locus with two alleles in a population in Hardy–Weinberg equilibrium.
With two alleles per locus the gene frequency that gives the maximum frequency of heterozygotes \( Q = 2pq \) is found when \( p = 0.5 \). Therefore, in F2 populations derived from elite × elite pure line crosses we expect maximum frequency of heterozygotes.

### 2.3 Means of Non-inbred Populations and Derived Families

A population of *phenotypes* (Fig. 2.2) can be characterized in terms of not only its gene and genotypic frequencies but also its mean and variance for a quantitative trait. Environmental factors largely influence the expression of these traits. These traits are studied by measures of central tendency and dispersion instead of phenotypic ratios. Genomic tools may provide additional information on gene information and the genetic architecture of quantitative traits as long as sample sizes are representative and a random set of populations is involved.

![Fig. 2.2](image.png) A population of phenotypes is made up of genotype and environment

A *phenotypic value* is an observed measure of its effect on the quantitative trait and can be measured. The values associated with genotypes are measured indirectly from the corresponding phenotypic values.

The phenotypic value can be divided into genotypic value and its environmental deviation as follows:

\[
P \text{ (phenotypic value)} = G \text{ (genotypic value)} + E \text{ (environmental deviation)}
\]

Therefore, phenotypic values are due to genetic and non-genetic circumstances. It is still challenging to accurately measure the *genotypic value* of an individual. However, it can be measured if we use a simple genetic model (one locus and two alleles) where the genotypes are distinguishable in their phenotype (e.g., inbred lines). So if we assign arbitrary values to the genotypes we can build a scale of genotypic values:

\[
\begin{array}{c|c|c}
A_2A_2 & 0 & A_1A_2 \\
-a & & d \\
A_1A_1 & a & \text{Mid-parent value}
\end{array}
\]

It depends on \( d/a \) or degree of dominance shown by the locus

where \( d \) and \( d/a \) are related to the level of dominance.

The degree of dominance for genes affecting plant height in maize is different from the level of dominance, if any, for genes affecting plant height in a self-pollinating crop like wheat. Hybrid vigor in maize is important and the difference
between inbred lines and hybrids for plant height is significant and so is the dominance level for genes affecting these types of traits.

How do gene frequencies affect the mean of a trait in a population?

Considering one locus with two alleles, \(A_1\) and \(A_2\), it is assumed that each locus has a particular effect on the total individual phenotype. Arbitrarily assuming \(A_1\) to be the allele that increases the value, we can denote by \(+a\), \(-a\), and \(d\) the effects of genotypes \(A_1A_1\), \(A_2A_2\), and \(A_1A_2\), respectively. Such effects are taken as deviations from the mean of the two homozygotes, as shown on the linear scale earlier.

The population mean is thus calculated considering both the genotypic frequencies and genotypic effects (coded values), as shown in Table 2.1. Let gene frequencies for ‘\(A_1\)’ and ‘\(A_2\)’ be \(p\) and \(q\), respectively.

**Table 2.1** Genotypic values and frequencies in a population in Hardy–Weinberg equilibrium for one locus with two alleles

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency ((F_i))</th>
<th># of ‘(A_1)’ alleles</th>
<th>Genotypic values (^a) ((X_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>(p^2)</td>
<td>2</td>
<td>(a) (d) (u) (z)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>(2pq)</td>
<td>1</td>
<td>(d) (h) (au) (\hat{h})</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>(q^2)</td>
<td>0</td>
<td>(-a) (-d) (-u) (o)</td>
</tr>
</tbody>
</table>

The first two columns show the three genotypes and their frequencies in a random mating population.

\(^a\)Different symbols across literature. We will use \(a\), \(d\), and \(-a\)

The mean value of a population is obtained by multiplying the values of each genotype \((X_i\) or genotypic effect) by their frequencies \((f_i)\). Then we sum over the three genotypes.

\[
\bar{X} = \sum (X_iF_i) \text{ or } \sum X_i/n
\]

Since the sum of frequencies is 1 \((p + q = 1)\) the sum of values multiplied by frequencies is the mean value:

\[
\bar{X} = p^2a + 2pqd - q^2a
\]

\[
(p^2 - q^2)a + 2pqd
\]

\[
(p+q)(p-q)a + 2pqd \quad \text{since } p + q = 1
\]

\[
(p-q)a + 2pqd
\]

\[
\bar{X} = \sum (p-q)a + 2\sum pqd
\]

After the contribution of several loci
As seen above, the mean will vary according to the level of dominance, the gene frequencies, and/or if genes become fixed. You could ask what would happen to the mean if \( d = 0 \), if \( A_1 \) was fixed, if \( d = a \), or if the population had frequencies in equilibrium. The contribution of any locus to the population mean has one term for homozygotes and another term for heterozygotes. The formula assumes that the combination of loci produces a joint additive effect on the trait. The ‘additive action’ is, therefore, associated not only with alleles at one locus but also with alleles at different loci. Alleles at a locus have additive action in the absence of dominance and across loci if epistatic deviations are not present. Since environmental effects are taken as deviations from the general mean over the whole population, they add to zero, and then also expresses the mean phenotypic value.

### 2.3.1 Half-Sib Family Means

A half-sib family is obtained from seeds produced by one plant (female common parent) that was pollinated by a random sample of pollen from the population (Table 2.2).

**Table 2.2** Genotypic values and frequencies of half-sib families from a population in Hardy–Weinberg equilibrium for one locus with two alleles

<table>
<thead>
<tr>
<th>Female parent</th>
<th>Frequency</th>
<th>Family genotypes(^a)</th>
<th>Coded half-sib family values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1A_1 )</td>
<td>( p^2 )</td>
<td>( p ) ( q ) ( - )</td>
<td>( pa + qd )</td>
</tr>
<tr>
<td>( A_1A_2 )</td>
<td>( 2pq )</td>
<td>( \frac{1}{2}p ) ( \frac{1}{2} ) ( \frac{1}{2}q )</td>
<td>( \frac{1}{2}[(p-q)a + d] )</td>
</tr>
<tr>
<td>( A_2A_2 )</td>
<td>( q^2 )</td>
<td>( - ) ( p ) ( q )</td>
<td>( pd - qa )</td>
</tr>
</tbody>
</table>

\(^a\)Produced after pollination by \( p(A_1) \) and \( q(A_2) \) male gametes.

The mean of the population of half-sib families is

\[
\bar{X}_{HS} = p^2 (pa - qd) + 2pq \left( \frac{1}{2} \right) \left[ (p - q) a + \left( \frac{1}{2} \right) d \right] + q^2 (pq - qa) \\
= (p - q)a + 2pqd
\]

This is equal to the original population mean.

### 2.3.2 Full-Sib Families

A full-sib family is obtained by crossing a random pair of plants (both parents in common) from the population. The probability of each cross is obtained by the product of genotypic frequencies, as shown in Table 2.3.
### 2.3 Means of Non-inbred Populations and Derived Families

#### Table 2.3 Genotypic values and frequencies of full-sib families from a population in Hardy–Weinberg equilibrium for one locus with two alleles

<table>
<thead>
<tr>
<th>Female parent</th>
<th>Male parent</th>
<th>Probability of cross</th>
<th>Family genotypes</th>
<th>Coded full-sib family values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$A_1A_1$</td>
<td>$p^4$</td>
<td>$A_1A_1$</td>
<td>$a$</td>
</tr>
<tr>
<td></td>
<td>$A_1A_2$</td>
<td>$2p^3q$</td>
<td>$\frac{1}{2}$ $\frac{1}{2}$ $-$</td>
<td>$(\frac{1}{2})(a + d)$</td>
</tr>
<tr>
<td></td>
<td>$A_2A_2$</td>
<td>$p^2q^2$</td>
<td>$-$ $1$ $-$</td>
<td>$(\frac{1}{2})(a + d)$</td>
</tr>
</tbody>
</table>

- The mean of the population for full-sib families is:
  \[
  \bar{X}_{FS} = p^4 (a) + 2p^3q (\frac{1}{2}) [(a + b) + \cdots + q^4] (-a) 
  = (p - q)a + 2pqd
  \]

Results so far obtained show that the expected value of half-sib families as well as of full-sib families equals the mean of the reference population.

#### 2.3.3 Inbred families

Selfing is the most common system of inbreeding used in practical maize breeding for inbred line development during pedigree selection. Considering a non-inbred parent population in Hardy–Weinberg equilibrium from which selfed lines will be drawn, we have the family structure as shown in Table 2.4 for $S_1$ families, i.e., families developed by one generation of selfing. This assumes that the $F_2$ population equals an $S_0$ and we will follow this nomenclature throughout the book. (Note that there are maize breeding programs assuming an $F_2$ population equals an $S_1$.)

#### Table 2.4 Genotypic values and frequencies of inbred ($S_1$) families from a non-inbred population in Hardy–Weinberg equilibrium for one locus with two alleles

<table>
<thead>
<tr>
<th>Parent genotypes</th>
<th>Frequency</th>
<th>Family genotypes</th>
<th>Coded $S_1$ family values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$p^2$</td>
<td>$A_1A_1$</td>
<td>$a$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$2pq$</td>
<td>$\frac{1}{4}$ $\frac{1}{2}$ $\frac{1}{4}$</td>
<td>$(\frac{1}{2})d$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$q^2$</td>
<td>$-$ $-$ $1$</td>
<td>$-a$</td>
</tr>
</tbody>
</table>

*aAfter one generation of selfing*
The mean of the population for inbred families is

$$\bar{X}_{S_1} = p^2a + 2pq \left( \frac{1}{2} \right) d + q^2 (-a)$$

$$= (p - q)a + 2pqd$$

which equals the reference population mean when $d = 0$, i.e., when there are no dominance effects. If dominance effects are present, the mean is reduced (see below). If the gene frequencies are the same for the reference and $S_1$ populations, the mean of the $S_1$ population will be halfway between the mean of the $S_0$ and $S_\infty$ generations.

Under a regular system of selfing the general mean decreases in each generation due to decreases in the frequency of heterozygotes. The general formula for the $n$th generation of inbreeding is

$$\bar{X}_{S_n} = (p - a)a + \left( \frac{1}{2} \right)^{n-1} pqd$$

which equals the non-inbred population mean for $n = 0$.

The above formula may also be expressed as a function of $F_n$, the coefficient of inbreeding, of progenies in the $n$th generation of selfing:

$$\bar{X}_{S_n} = (p - q)a + 2 \left( 1 - F_n \right) pqd$$

which equals the non-inbred population mean when $F = 0$.

Most reference or base populations (first segregating population, $S_0$ in maize) is derived from elite (pure line) \times elite (pure line) crosses. Therefore, average gene frequency at all segregating loci is expected to be $\frac{1}{2}$ and, therefore, we may assume that $F_2$ populations can be represented with loci having gene frequencies in equilibrium ($p = q = \frac{1}{2}$). Linkage, however, could be a serious bias. For example, most commercial breeding programs are represented within this scheme. However, the maize-breeding program at NDSU also develops inbred lines from genetically broad-based populations with arbitrary allele frequencies.

If we cross two inbred lines and consider one locus (two alleles A and a in this case):

<table>
<thead>
<tr>
<th>Parents</th>
<th>AA</th>
<th>×</th>
<th>Aa</th>
<th>⊗</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>F_1</td>
<td>AA</td>
<td></td>
<td>Aa</td>
<td>⊗</td>
<td>aa</td>
</tr>
</tbody>
</table>

The $S_0$ population is considered to be the base population:

$$F_2 (S_0) \quad \left( \frac{1}{4} \right) AA \quad \left( \frac{1}{2} \right) Aa \quad \left( \frac{1}{4} \right) aa$$

Base population or reference population (alleles segregating), e.g., 200 individuals

The $S_1$ population is considered to be the result of one generation of self-fertilization:

$$F_3 (S_1) \quad \left( \frac{1}{4} \right) AA \quad \left( \frac{1}{2} \right) \left[ \left( \frac{1}{4} \right) AA + \left( \frac{1}{2} \right) Aa + \left( \frac{1}{4} \right) aa \right] \quad \left( \frac{1}{4} \right) aa$$
2.4 Means of Inbred Populations and Derived Families

So, if we calculate the mean for each generation we obtain

\[ \bar{X}_{S_0} = (\frac{1}{4}) AA + (\frac{1}{2}) Aa + (\frac{1}{4}) aa \]
\[ = (\frac{1}{4}) a + (\frac{1}{2}) d - (\frac{1}{4}) a \]
\[ = (\frac{1}{2}) d \]
\[ \bar{X}_{S_1} = (\frac{1}{4}) AA + (\frac{1}{2}) [(\frac{1}{4}) AA + (\frac{1}{2}) Aa + (\frac{1}{4}) aa] + (\frac{1}{4}) aa \]
\[ = (\frac{1}{4}) a + (\frac{1}{2}) [(\frac{1}{2}) d] - (\frac{1}{4}) a \]
\[ = (\frac{1}{4}) d \]

So, if we come back to our scale of genotypic values we see that the mean is reduced in the presence of dominance effects when selfing:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Coded genotypic values</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>0</td>
<td>aa</td>
</tr>
<tr>
<td>Aa</td>
<td>(\frac{1}{4}) d</td>
<td>a</td>
</tr>
<tr>
<td>AA</td>
<td>(\frac{1}{2}) d</td>
<td>a</td>
</tr>
</tbody>
</table>

The examples of plant height and grain yield in maize seem to closely validate the theory.

2.4 Means of Inbred Populations and Derived Families

The main difference between inbred and non-inbred populations is in genotypic frequencies. Gene frequencies remain constant, but genotypic frequencies change under inbreeding because inbreeding decreases the frequency of heterozygotes and consequently increases the frequency of homozygous genotypes. Using Wright’s coefficient \( F \) as a measure of inbreeding, genotypic frequencies are distributed according to the pattern shown in Table 2.5.

Table 2.5 Genotypic values and frequencies in inbred populations (inbreeding measured by \( F \)) for one locus with two alleles

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequencies</th>
<th>Coded genotypic values</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1A_1</td>
<td>2p^2 + Fpq</td>
<td>a</td>
</tr>
<tr>
<td>A_1A_2</td>
<td>2pq(1-F)</td>
<td>d</td>
</tr>
<tr>
<td>A_2A_2</td>
<td>q^2 + Fpq</td>
<td>-a</td>
</tr>
</tbody>
</table>

Hence the mean of an inbred population is

\[ \bar{X}_S = (p - q)a + 2pq(1 - F)d \]

When \( F = 1 \) (completely homozygous population), then the inbred population mean equals \((p - q)a \) because there will be no dominance effects expressed.
When \( F = \frac{1}{2} \), then the inbred population mean becomes \((p - q)a + pqd\), which equals the \( S_1 \) family mean, as previously shown in Table 2.4.

Half-sib and full-sib families drawn from an inbred population result in non-inbred progenies, and their mean equals that of a non-inbred population \((p - q)a + 2pqd\) because one generation of random mating is involved.

Kempthorne (1957) gives a general formulation for the changes in population mean under inbreeding, including epistatic effects. In his definition the lack of dominance and dominance types of epistasis do not change the population mean with inbreeding. If there are no dominance types of epistasis, the mean of the inbred population is linearly related to \( F \) even in the presence of additive types of epistasis.

### 2.5 Mean of a Cross Between Two Populations

Let \( P_1 \) and \( P_2 \) be two populations in Hardy–Weinberg equilibrium. Denoting by \( p \) and \( q \) the frequencies of both alleles, \( A_1 \) and \( A_2 \), in population \( P_1 \) and by \( r \) and \( s \) the frequencies of the same alleles in population \( P_2 \), we have the following structure in the crossed population (Table 2.6).

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency</th>
<th>Coded genotypic values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1A_1 )</td>
<td>( pr )</td>
<td>( a )</td>
</tr>
<tr>
<td>( A_1A_2 )</td>
<td>( ps + qr )</td>
<td>( d )</td>
</tr>
<tr>
<td>( A_2A_2 )</td>
<td>( qs )</td>
<td>(-a)</td>
</tr>
</tbody>
</table>

The population cross mean for one locus is

\[
\overline{X}_{12} = (pr - qs)a + (ps + qr)d
\]

The cross between two populations also may be obtained according to a family structure. If half-sib families are drawn with, for example, \( P_1 \) as female parents, we have the family structure shown in Table 2.7.

The mean of half-sib families is then

\[
\overline{X}_{HS_{12}} = p^2(ra + sd) + 2pq \left[ \left( \frac{1}{2} \right)(r - s)a + \left( \frac{1}{2} \right)d \right] + q^2(rd - sa)
\]

which equals the randomly crossed population mean. Note that the mean will be the same whatever population is used as the female parent.

If the crossed population is structured as full-sib families, we have the genotypes and frequencies shown in Table 2.8.
Table 2.7 Genotypic values and frequencies in a cross between two populations structured as half-sib families for one locus with two alleles

<table>
<thead>
<tr>
<th>Female parent, ( P_1 )</th>
<th>Frequencies</th>
<th>Family genotypes(^a)</th>
<th>Coded half-sib family values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1 A_1 )</td>
<td>( p^2 )</td>
<td>( r ) ( s )</td>
<td>( ra + sd )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( 2pq )</td>
<td>( (\frac{1}{2})r ) ( (\frac{1}{2})(r + s) ) ( (\frac{1}{2})s )</td>
<td>( (\frac{1}{2})(r - s)a + (\frac{1}{2})d )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( q^2 )</td>
<td>( r ) ( s )</td>
<td>( rd - sa )</td>
</tr>
</tbody>
</table>

\(^a\)After pollination by \( r(A_1) \) and \( s(A_2) \) male gametes (from \( P_2 \))

Table 2.8 Genotypic values and frequencies in a cross between two populations structured as full-sib families for one locus with two alleles

<table>
<thead>
<tr>
<th>Female parent, ( P_1 )</th>
<th>Male parent, ( P_2 )</th>
<th>Frequency of crosses</th>
<th>Family genotypes</th>
<th>Coded full-sib family values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1 A_1 )</td>
<td>( A_1 A_1 )</td>
<td>( p^2 r^2 )</td>
<td>( 1 ) ( 0 ) ( 0 )</td>
<td>( a )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( A_1 A_2 )</td>
<td>( 2p^2 rs )</td>
<td>( \frac{1}{2} ) ( \frac{1}{2} ) ( 0 )</td>
<td>( (\frac{1}{2})(a + d) )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( A_2 A_2 )</td>
<td>( p^2 s^2 )</td>
<td>( 0 ) ( 1 ) ( 0 )</td>
<td>( d )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( A_1 A_1 )</td>
<td>( 2pqr^2 )</td>
<td>( \frac{1}{2} ) ( \frac{1}{2} ) ( 0 )</td>
<td>( (\frac{1}{2})(a + d) )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( A_1 A_2 )</td>
<td>( 4pqrs )</td>
<td>( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} )</td>
<td>( (\frac{1}{2})d )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( A_2 A_2 )</td>
<td>( 2pqrs^2 )</td>
<td>( 0 ) ( \frac{1}{2} ) ( \frac{1}{2} )</td>
<td>( (\frac{1}{2})(d - a) )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( A_2 A_2 )</td>
<td>( q^2 r^2 )</td>
<td>( 0 ) ( 1 ) ( 0 )</td>
<td>( d )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( A_1 A_2 )</td>
<td>( 2q^2 rs )</td>
<td>( 0 ) ( \frac{1}{2} ) ( \frac{1}{2} )</td>
<td>( (\frac{1}{2})(d - a) )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( A_2 A_2 )</td>
<td>( q^2 s^2 )</td>
<td>( 0 ) ( 0 ) ( 1 )</td>
<td>( -a )</td>
</tr>
</tbody>
</table>

The mean becomes

\[
\bar{X}_{FS12} = p^2 r^2 a + 2p^2 rs (\frac{1}{2})(a + d) + \cdots + q^2 s^2 (-a) = (pr - qs)a + (ps + qr)d
\]

which again equals the randomly crossed population mean and has the same value whatever parent is used as female.

2.6 Average Effect

It is a value not associated with genotypes but rather associated with the genes carried by the individual and transmitted to its offspring. The average effect of an allele is the mean deviation from the population mean of individuals (Table 2.9).

Table 2.9 Genotypic values, frequencies, and average effect of alleles in a one locus model

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency ((F_i))</th>
<th>Genotypic values ((Y_i))</th>
<th>‘(A_1)’ gametes</th>
<th>‘(A_2)’ gametes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1 A_1 )</td>
<td>( p^2 )</td>
<td>( a )</td>
<td>( p )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( A_1 A_2 )</td>
<td>( 2pq )</td>
<td>( d )</td>
<td>( q )</td>
<td>( p )</td>
</tr>
<tr>
<td>( A_2 A_2 )</td>
<td>( q^2 )</td>
<td>( -a )</td>
<td>( 0 )</td>
<td>( q )</td>
</tr>
</tbody>
</table>
So, the average effect of ‘A1’ alleles ($\alpha_1$) is

$$\alpha_1 = pa + qd - \bar{X}$$

$$= pa + qd - [(p - q)a + 2pqd]$$

$$\alpha_1 = q[a + (q - p)d]$$

and the average effect of ‘A2’ alleles ($\alpha_2$) is

$$\alpha_2 = pd + qa - \bar{X}$$

$$= pd - qa - [(p - q)a + 2pqd]$$

$$\alpha_2 = -p[a + (q - p)d]$$

The concept of ‘average effect’ of a gene is basic to the understanding of breeding value. The average effect of a gene is defined as the mean deviation from the population mean of a group of individuals that received the gene from the same parent, the other gene of such individuals being randomly sampled from the whole population as shown in Table 2.10.

**Table 2.10** Genotypes and average effects of progenies having a common parental gamete for one locus (after Falconer, 1960)

<table>
<thead>
<tr>
<th>Gamete</th>
<th>Genotypes in progenies</th>
<th>Progeny effects</th>
<th>Population mean</th>
<th>Average effect of a gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>$A_1A_1$ $A_1A_2$ $A_2A_2$</td>
<td>$pa + qd$</td>
<td>$(p - q)a + 2pqd$</td>
<td>$\alpha_1 = q[a + (q - p)d]$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>$A_2A_1$ $A_1A_2$ $A_2A_2$</td>
<td>$pd - qa$</td>
<td>$(p - q)a + 2pqd$</td>
<td>$\alpha_2 = -p[a + (q - p)d]$</td>
</tr>
</tbody>
</table>

*After pollination with a random sample of gametes: $p(A_1)$ and $q(A_2)$

If two alleles are present per locus we can define the *average effect of gene substitution* ($\alpha$) as the difference between the average effects of the two alleles:

$$\alpha = \alpha_1 - \alpha_2$$

$$= p + q[a + (q - p)d]$$

$$\alpha = a + (q - p)d$$

The ‘average effect of a gene substitution’ is the average deviation due to the substitution of one gene by its allele in each genotype. Consider the $A_2$ gene being substituted by its gene $A_1$ at random in the population. Since the $A_1A_1$, $A_1A_2$, and $A_2A_2$ genotypes have frequencies $p^2$, $2pq$, and $q^2$, respectively, then genes that will be substituted are found in genotypes $A_1A_2$ and $A_2A_2$ with frequency $pq + q^2 = q$. Proportionally, we have $pq/q = p A_1A_2$ genotypes; $q^2/lq = q A_2A_2$ genotypes; i.e., $A_2$ genes will be substituted in $A_1A_2$ and $A_2A_2$ genotypes at frequencies $p$ and $q$, respectively.
respectively. When the substitution is in $A_1A_2$ genotypes, the change in genotypic value will be from $d$ to $a$, and when substitution takes place in $A_2A_2$, the change is from $-a$ to $d$. The change in the population is

$$\alpha = p(a - d) + q(d + a)$$

$$= a + (q - p)d$$

This is the definition of the average effect of a gene substitution. It can be seen that the average effect of a gene substitution $\alpha$ is the difference between the average effects of genes involved in the substitution; i.e., $\alpha = \alpha_1 - \alpha_2$, as shown in Table 2.10. Both the average effect of a gene and the average effect of a gene substitution depend on gene effects and gene frequency; therefore, both are a property of the population and of the gene.

### 2.7 Breeding Value

When panmictic populations are under consideration, one must consider that the genotypes of any offspring are not identical to their parents. The relationship between any individual in the offspring and one of its parents is established by the gamete received from that parent. It is known that gametes are haploid entities and carry genes and not genotypes. So for the understanding of the inheritance of a quantitative trait in a panmictic population it is valuable to have an individual measure associated with its genes and not its genotype. Such a value is designated by Falconer (1960) as ‘breeding value,’ which is ‘the value of an individual, judged by the mean value of its progeny.’

The breeding value of an individual can, therefore, be measured. It is twice the mean deviation of the progeny from the population mean since only one half of the genes are passed to the progeny.

The breeding value of an individual is equal to the sum of average effects of the genes it carries (Table 2.11). If all loci are taken into account, the breeding value of a particular genotype is the sum of breeding values from each locus (‘additive genotype’).

Extending the concept of average effect of genes to the individual genotype gives the concept of breeding value of the individual. At the gene level the breeding value is the sum of average effects of genes, summation being over all alleles and over all loci. Similar to the average effect of a gene, breeding value is a property of the

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency</th>
<th>Genotypic values</th>
<th>Breeding value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$p^2$</td>
<td>$a$</td>
<td>$2\alpha_1$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$2pq$</td>
<td>$d$</td>
<td>$\alpha_1 + \alpha_2$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$q^2$</td>
<td>$-a$</td>
<td>$2\alpha_2$</td>
</tr>
</tbody>
</table>
individual as well as of the population but the breeding value can be measured experimentally. The breeding value is, therefore, a measurable quantity and is of much relevance in animal breeding where the individual value is an important criterion. On the other hand, individual values are less important in crop species like maize, since the whole population is concerned. In this case individuals are looked upon as ephemeral representatives of the whole population and its gene pool. The average effect of a gene and individual breeding value concepts, however, are closely related to genotype evaluation procedures like topcross tests in maize.

### 2.8 Genetic Variance

Breeders choose not only populations with high phenotypic means but also populations having large and useful genetic variance.

The variation among phenotypic values (phenotypic variance) can be partitioned into observational components of variance:

$$\hat{\sigma}_P^2 = \hat{\sigma}_G^2 + \hat{\sigma}_E^2 + \hat{\sigma}_{GE}^2$$

Even though it is the goal of breeders to separate the genetic variance ($\hat{\sigma}_G^2$) from the environmental variance ($\hat{\sigma}_E^2$), the variance due to crossover (e.g., rank) and non-crossover (e.g., magnitude) interactions between genotypes and environments ($\hat{\sigma}_{GE}^2$) is the most difficult to manage.

Fisher (1918) first demonstrated that the hereditary variance in a random mating population can be partitioned into three parts: (1) an additive portion associated with average effects of genes, (2) a dominance portion due to allelic interactions, and (3) a portion due to non-allelic interactions or epistatic effects. Therefore, the genetic proportion of variance has the following components:

$$\hat{\sigma}_G^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_I^2$$

**Epistatic** interactions give rise to the component of variance $\hat{\sigma}_I^2$, which is the variance due to the interaction deviations involving more than one locus. This is subdivided into components according to the number of loci involved (e.g., two-factor interaction, three-factor interaction). Another subdivision can be done based upon the type of interaction present. If the interaction involves breeding values then the additive $\times$ additive interaction variance is present ($\hat{\sigma}_{AA}^2$). If the interaction is between the breeding value of one locus and the dominance deviation of the other then the additive $\times$ dominance interaction variance is present ($\hat{\sigma}_{AD}^2$). Finally, if the interaction is between dominance deviations from two loci then the dominance $\times$ dominance interaction variance is present ($\hat{\sigma}_{DD}^2$).

A general theory for the partition of hereditary variance was further developed by Cockerham (1954) and Kempthorne (1954). Thus, in general, the total genetic variance $\hat{\sigma}_G^2$ can be partitioned into the following components:
2.8 Genetic Variance

\[ \hat{\sigma}_A^2 \] additive variance due to the average effects of alleles (additive effects, same locus)

\[ \hat{\sigma}_D^2 \] dominance variance due to interaction of average effects of alleles (dominance effects, same locus)

\[ \hat{\sigma}_{AA}, \hat{\sigma}_{AAA}, \ldots = \text{epistatic variances due to interaction of additive effects of two or more loci} \]

\[ \hat{\sigma}_{DD}, \hat{\sigma}_{DDD}, \ldots = \text{epistatic variances due to interaction of dominance effects of two or more loci} \]

\[ \hat{\sigma}_{AD}, \hat{\sigma}_{AAD}, \hat{\sigma}_{ADD}, \ldots = \text{epistatic variances due to interaction of additive and dominance effects involving two or more loci} \]

Collecting all components together, the total genetic variance is

\[ \hat{\sigma}_G^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{DD}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{AAA}^2 + \hat{\sigma}_{AAD}^2 + \cdots \]

The interaction between loci (located within or between chromosomes) controlling the expression of quantitative traits is assumed to be frequent. However, the estimation of the amount of variance generated by interactions is challenging even at the molecular level.

An additional source of genetic variance is the one due to disequilibrium. In this case, genotypic frequencies at several loci cannot be predicted by allele frequencies. If there is no epistasis between two loci we can estimate the total genotypic variance caused by the two loci together as follows:

\[ \hat{\sigma}_{TG}^2 = \hat{\sigma}_G^2 \text{ (first locus)} + \hat{\sigma}_D^2 \text{ (second locus)} + 2\text{Cov} \text{ (both loci)} \]

The covariance term is the correlation between the genotypic values at the two loci in different individuals. This correlation can be positive or negative. Therefore, linkage disequilibrium can either decrease or increase the variance depending on the linkage phase present. Coupling phase linkage will cause an upward bias for the additive and dominance genetic variances. On the other hand, repulsion phase linkage will only cause the dominance genetic variance to increase; the additive genetic variance is expected to decrease. No covariance term is present if there is random mating equilibrium.

Linkage of traits with molecular markers became a popular scientific research targeted initially at improvement of quantitative traits. But quantitative traits are dependent upon a large number of genes each having a relatively minor effect as compared with environmental effects (Lonnquist, 1963). Based on this definition, quantitative traits have been explained by polygenes (Mather, 1941) and quantitative trait loci (QTL) (Geldermann, 1975) or chromosome segments affecting the quantitative trait (Falconer and Mackay, 1996). Rather than QTL mapping in bi-parental populations breeding plans that currently utilize molecular marker information for germplasm (e.g., association mapping on relevant breeding germplasm, genome-wide selection) could be assessed depending on the amount of linkage between markers and loci and may generate useful information that is relevant to improving elite germplasm.
(Sorrells, 2008). All these approaches, however, rely on the maintenance of strong applied breeding programs and need to be proven useful for developing cultivars in a more efficient way. Alternative approaches such as ‘meta-QTL analysis’ focused on major QTLs that are stable across numerous populations also have potential (Snape et al., 2008).

### 2.8.1 Total Genetic Variance

Total genetic variance of a population in Hardy–Weinberg equilibrium is obtained from a modified version of Table 2.1 as follows:

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency ((F_i))</th>
<th>Genotypic values ((GV))</th>
<th>(F_i \times GV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>(p^2)</td>
<td>(a)</td>
<td>(p^2a)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>(2pq)</td>
<td>(d)</td>
<td>(2pqd)</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>(q^2)</td>
<td>(-a)</td>
<td>(q^2(-a))</td>
</tr>
</tbody>
</table>

We could, therefore, estimate the variance statistically and use this information to estimate the genetic variance of a population:

\[
\hat{\sigma}^2_G = \sum F_i X_i^2 - \sum(F_i X_i)^2 \text{ or } \sum X_i^2 - \frac{(\sum X_i)^2}{n} \text{ or } \sum(X_i - X)^2
\]

In this case, the mean is the corrector factor in the working formula. Therefore,

\[
\hat{\sigma}^2_G = [p^2a^2 + 2pqd^2 + q^2(-a)^2] - \bar{X}^2 \quad \text{and} \quad \bar{X}^2 = [(p-q)^2a^2 + 4pq(p-q)ad + 4p^2q^2d^2]
\]

Then, the total genetic variance for one locus is

\[
\hat{\sigma}^2_G = p^2a^2 + 2pqd^2 + q^2(-a)^2 - p^2a^2 + 2pqd^2 - 4pq(p-q)ad - 4p^2q^2d^2
\]

Then, the total genetic variance for one locus is

\[
\hat{\sigma}^2_G = 2pq[a^2 + b^2 - 2(p-q)ad - 2pqd^2]
\]

And if we have a population with frequencies in equilibrium \((p = \frac{1}{2} \text{ and } q = \frac{1}{2})\) then (e.g., \(F_2\) populations)
2.8 Genetic Variance

\[ \hat{\sigma}^2_G = \left(\frac{1}{2}\right) a^2 + \left(\frac{1}{4}\right) d^2 \]

The first term of the formula is the most important for breeders while the second term is the one breeders are not able to fix.

### 2.8.2 Additive Genetic Variance

The additive genetic variance is obtained as follows:

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency ((F_i))</th>
<th>Breeding values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>(p^2)</td>
<td>(2\alpha_1 = 2qa)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>(2pq)</td>
<td>(\alpha_1 + \alpha_2 = (q-p)a)</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>(q^2)</td>
<td>(2\alpha_2 = -2p\alpha)</td>
</tr>
</tbody>
</table>

The additive genetic variance \(\hat{\sigma}^2_A\) is, therefore derived as follows:

\[
\hat{\sigma}^2_A = p^2 (2qa)^2 + 2pq [(q-p)a]^2 + q^2 [-2p\alpha]^2
\]

\[
\hat{\sigma}^2_A = 2pq\alpha^2 [2pq + q^2 + p^2]
\]

Then, the additive genetic variance for one locus is

\[
\hat{\sigma}^2_A = 2pq\alpha^2 \quad \text{or} \quad 2pq \left[ a + (q-pa) \right]^2
\]

Clearly, \(\hat{\sigma}^2_A\) depends on gene frequencies. Therefore, for segregating alleles in equilibrium (e.g., \(p = q = 0.5\)) then

\[
\hat{\sigma}^2_A = 2pq \left[ a + (q-p) \right] \left[ a + (q-p) \right]^2
\]

\[
= 2pq a^2
\]

\[
\hat{\sigma}^2_A = \left(\frac{1}{2}\right) a^2
\]

This is the additive genetic variance for the special case of \(F_2\) populations.

In the general case of arbitrary gene frequencies (e.g., genetically broad-based populations) the following formula applies:

\[
\hat{\sigma}^2_s = \hat{\sigma}^2_A
\]

Note the \(\hat{\sigma}^2_A\) has a level of dominance between alleles in the additive portion of this segregating population.
2.8.3 Dominance Genetic Deviations

The dominance variance is the remainder from the total variance and is calculated by subtraction as

\[
\hat{\sigma}_D^2 = \hat{\sigma}_G^2 - \hat{\sigma}_A^2 \\
= 2pq \left[ a^2 + 2(q - p) ad + (1 - 2pq) d^2 \right] - 2pq \left[ a + (q - p) d \right]^2 \\
= 2pq d^2 (2pq)
\]

Then, the dominance genetic variance for one locus is

\[
\hat{\sigma}_D^2 = 4p^2 q^2 d^2
\]

\(\hat{\sigma}_D^2\) also depends on gene frequencies. Therefore, for segregating alleles in equilibrium (e.g., \(p = q = 0.5\)) then

\[
\hat{\sigma}_D^2 = 4(\tfrac{1}{4})(\tfrac{1}{4}) d^2 \\
\hat{\sigma}_D^2 = (\tfrac{1}{4}) d^2
\]

This is the dominance genetic variance for the special case of F2 populations.

In the general case of arbitrary gene frequencies (e.g., genetically broad-based populations) the following formula applies:

\[
\hat{\sigma}_D^2 = \hat{\sigma}_G^2 - \hat{\sigma}_A^2
\]

Note the \(\hat{\sigma}_D^2\) does not have additive effects in the dominance portion of this segregating population

Both additive and dominance genetic variances (\(\hat{\sigma}_A^2\) and \(\hat{\sigma}_D^2\)) can also be explained by regression analyses (Falconer and Mackay, 1996). \(\hat{\sigma}_A^2\) is defined as the variance due to linear regression of genotypic values on gene frequencies in individual genotypes explained by the sum squares of the regression ANOVA while \(\hat{\sigma}_D^2\) represents the variance due to deviations from the regression. Hence the dominance variance is the deviation from the regression of genotypic values on the gene content of the genotypes.

Therefore, if we look at the example for the additive genetic variance we obtain

\[
\hat{\sigma}_A^2 = \left[ \frac{\text{Cov}}{\text{Var}} \right]^2 \\
= \frac{(2pq \left[ a + (q - p) d \right])^2}{2pq}
\]

\[
\hat{\sigma}_A^2 = 2pq \left[ a + (q - p) d \right]^2
\]
Table 2.12 Gene frequency ($p$) at maximum values for $\hat{\sigma}_G^2$, $\hat{\sigma}_A^2$, and $\hat{\sigma}_D^2$ for no dominance and complete dominance for one locus with two alleles

<table>
<thead>
<tr>
<th>Gene action</th>
<th>$\hat{\sigma}_G^2$</th>
<th>$\hat{\sigma}_A^2$</th>
<th>$\hat{\sigma}_D^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No dominance</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>—</td>
</tr>
<tr>
<td>Complete dominance</td>
<td>$1 - \sqrt{\frac{1}{2}}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

Gene frequencies, which give the maximum values for $\hat{\sigma}_G^2$, $\hat{\sigma}_A^2$, and $\hat{\sigma}_D^2$, are found by taking the first derivative of their respective expressions equal to zero. The simplest cases are those for no dominance and complete dominance, as shown in Table 2.12.

2.8.4 Variance Among Non-inbred Families

Genetic variance among half-sib families from a population in Hardy–Weinberg equilibrium is obtained from Table 2.2 as follows:

$$\hat{\sigma}_{HS}^2 = \sum F_i X_i^2 - \bar{X}_{HS}^2 = \left(\frac{1}{2}\right) pq \left[a + (1 - 2p)d\right]^2$$

Since,

$$\hat{\sigma}_A^2 = 2pq \left[a + (q - p)d\right]^2$$

then,

$$\hat{\sigma}_{HS}^2 = \left(\frac{1}{4}\right) \hat{\sigma}_A^2$$

Theoretically, the total genetic variance (among and within half-sib families) equals the total genetic variance in the reference population, i.e., $\hat{\sigma}_G^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2$. From this total variance ($\frac{1}{4}$), $\hat{\sigma}_A^2$ is expressed among half-sib families.

The remainder, ($\frac{3}{4}$) $\hat{\sigma}_A^2 + \hat{\sigma}_D^2$, is expected to be present within half-sib families over the entire population of families.

Genetic variance among full-sib families is obtained from Table 2.3 as

$$\hat{\sigma}_{FS}^2 = \sum F_i X_i^2 - \bar{X}_{FS}^2 = pq \left[a + (q - p)d\right]^2 + p^2 q^2 d^2$$

Since,

$$\hat{\sigma}_D^2 = 4p^2 q^2 d^2$$
then,

\[ \hat{\sigma}^2_{FS} = (\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D. \]

Relative to the total genetic variance of a population it is seen that \(\hat{\sigma}^2_{FS} = (\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D\). This is the portion of the total genetic variance expressed among families. Thus the remainder of the total genetic variance, \((\frac{1}{2})\hat{\sigma}^2_A + (\frac{3}{4})\hat{\sigma}^2_D\), is expected to be present within families over the whole population.

If epistasis is considered, the approximation to the real genetic variance is \(\hat{\sigma}^2_G = \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_{AA} + \hat{\sigma}^2_{DD} + \hat{\sigma}^2_{AAD} + \cdots\), where all components were defined previously.

In the same way, \(\hat{\sigma}^2_{HS} = (\frac{1}{4})\hat{\sigma}^2_A + (\frac{1}{16})\hat{\sigma}^2_{AA} + \cdots\), or simply \(\hat{\sigma}^2_{HS} = (\frac{1}{4})\hat{\sigma}^2_A + \cdots\), indicating a bias due to epistatic components of variance. The variance among full-sib families is \(\hat{\sigma}^2_{FS} = (\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D + (\frac{1}{4})\hat{\sigma}^2_{AA} + \cdots\), or simply \(\hat{\sigma}^2_{FS} = (\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D + \cdots\), indicating a bias due to epistatic components.

The relation between the variance among families and the covariance between relatives within families are presented in Chapter 3.

### 2.8.5 Variance Among Inbred Families

Variance among \(S_1\) families from a non-inbred reference population is obtained from Table 2.4 as

\[
\hat{\sigma}^2_{s1} = \sum F_i X_i^2 - \bar{X}_{s1} = 2pq[a + (\frac{1}{2}) (q - p) d]^2 + p^2 q^2 d^2
\]

A problem that arises with inbreeding is that genetic variance among inbred families is not linearly related to genetic components of variance of the reference population. From the above expression it can be seen that \(\hat{\sigma}^2_{s1}\) can be translated into \(\hat{\sigma}^2_A\) and/or \(\hat{\sigma}^2_D\) only in the following situations:

- No dominance (\(d_i = 0\) for all loci): \(\hat{\sigma}^2_{s1} = \hat{\sigma}^2_A\)
- Gene frequency \(\frac{1}{2}\) for all loci: \(\hat{\sigma}^2_{s1} = \hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D\)

If the restriction of no dominance is imposed, it can also be demonstrated that the genetic variance among \(S_2\) families (after two generations of selfing) is \(\hat{\sigma}^2_{s2} = (\frac{3}{2})\hat{\sigma}^2_A\). In the same way, if gene frequencies are assumed to be \(\frac{1}{2}\), then \(\hat{\sigma}^2_{s2} = (\frac{3}{2})\hat{\sigma}^2_A + (\frac{3}{16})\hat{\sigma}^2_D\).

An alternative way to understand what was explained above is by emphasizing there are two types of variances: among and within progenies. The one among progenies does not generate additive values from heterozygote plants since the variation
among S1 plants within progenies is not included. In addition, the variance within parenthesis resembles to the variance among F2 individuals.

If the genetic variance for an F2 generation is

\[
\hat{\sigma}_2^2 = \sum g_i^2 - (\bar{X})^2 \\
= [(1/4) a^2 + (1/2) d^2] - [(1/2) d]^2 \\
= (1/2) a^2 + (1/2) d^2 - (1/4) d^2 \\
= (1/2) a^2 + (1/4) d^2
\]

\[
\hat{\sigma}_F^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 
\]

Variance among individuals (individual plant basis)

The variance among S1 families follows:

\[
\hat{\sigma}_S^2 = [(1/4) a^2 + (1/2) (0 + (1/4) a^2) + (1/4) d^2] - [(1/4) d]^2 \\
= [(1/2) a^2 + (1/2) (1/4) d^2] - (1/16) d^2 \\
= (1/2) a^2 + (1/16) d^2
\]

\[
\hat{\sigma}_S^2 = \hat{\sigma}_A^2 + (1/4) \hat{\sigma}_D^2
\]

While the variance within S1 families is

\[
\hat{\sigma}_w^2 = [0 + (1/2)(1/4) a^2 + (1/4) d^2 + 0] \\
= (1/4) a^2 + (1/8) d^2
\]

\[
\hat{\sigma}_w^2 = (1/2) \hat{\sigma}_A^2 + (1/2) \hat{\sigma}_D^2
\]

More details can be found in Section 4.12.

### 2.8.5.1 Distribution of Genetic Variances Among and Within Lines

If we continue selfing from the F3(S1) generation to the F8(S6) generation of inbreeding the distribution of variances among and within lines changes depending on the inbreeding coefficient \( F = 1 - (1/2)^n \) being ‘n’ the number of generations of continuous selfing (Table 2.13).

Assuming \( p = q = 1/2 \) then at the S6 level of inbreeding \( F \sim 1 \) \( \hat{\sigma}_G^2 = 2 \hat{\sigma}_A^2 \) among lines which have important breeding implications. As the inbreeding coefficient approaches unity the additive genetic variance among lines approaches twice the additive genetic variance of the reference population without inbreeding. Therefore,
Table 2.13  Distribution of variances among and within lines under continuous selfing assuming \( p = q = 0.5 \) and \( F = 1 - \left( \frac{1}{2} \right)^n \)

| Generation | \( F \) | Among lines | | Within lines | | Total |
|------------|-------|-------------|-----------------|-----------------|-----------------|
| S1         | 1/2   | 1           | 1/4             | 1/2             | 3/2             | 3/4             |
| S2         | 3/4   | 3/2         | 3/16            | 1/4             | 7/4             | 7/16            |
| S3         | 7/8   | 7/4         | 7/64            | 1/8             | 15/8            | 15/64           |
| S4         | 15/16 | 15/8        | 15/256          | 1/16            | 31/16           | 31/256          |
| S5         | 31/32 | 31/16       | 31/1024         | 1/32            | 63/32           | 63/1024         |
| S6         | 63/64 | 63/32       | 63/4096         | 1/64            | 127/64          | 127/4096        |
| \( \infty \) | 1     | 2           | 0               | 0               | 2               | 0               |

Differences among lines increases and breeders will be more effective in identifying the best lines than to identify the best parents among plants that are not inbred.

A summary of the distribution of \( \hat{\sigma}^2_A \) and \( \hat{\sigma}^2_D \) among and within lines for successive generations of inbreeding is given in Table 2.13. Two important features are that (1) the genetic variance among inbred families increases and within inbred families decreases with increased inbreeding and (2) the total genetic variance doubles from \( F = 0 \) to \( F = 1 \), and all the genetic variance at \( F = 1 \) is additive. General limitations when inbreeding is involved are discussed in Chapter 3.

### 2.8.6 Variance in Inbred Populations

The effect of inbreeding is to increase total genetic variance in the population, and such an increase depends on the level of inbreeding. Using \( u_s \) as an alternative symbol for the mean of the inbred population the total genetic variance is calculated from Table 2.5 as follows:

\[
\hat{\sigma}^2_{GS} = (p^2 + Fpq)a^2 + 2pq(1 - F)d^2 + (q^2 + Fpq)a^2 - u_s^2
\]

\[
= 2pq(1 + F) \left[ a + \frac{1-F}{1+F}(q-p) \right]^2 + 4pq \frac{1-F}{1+F}(p+Fq)(q+Fp)d^2
\]

which equals the genetic variance of a non-inbred population when \( F = 0 \). The additive genetic variance is

\[
\hat{\sigma}^2_{AS} = 2pq(1 + F) \left[ a + \frac{1-F}{1+F}(q-p) \right]^2
\]

and the dominance variance is again calculated as \( \hat{\sigma}^2_{GS} - \hat{\sigma}^2_{AS} \):

\[
\hat{\sigma}^2_{DS} = 4pq \frac{1-F}{1+F}(p+Fq)(q+Fp)d^2
\]
When \( F = 0 \), \( \hat{\sigma}_{AS}^2 = \hat{\sigma}_A^2 \) and \( \hat{\sigma}_{DS}^2 = \hat{\sigma}_D^2 \), respectively. For \( F > 0 \), \( \hat{\sigma}_{AS}^2 \) and \( \hat{\sigma}_{DS}^2 \) cannot be expressed by \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) nor be translated from one generation of inbreeding to another. If gene frequency is assumed to be \( \frac{1}{2} \), then \( \hat{\sigma}_{AS}^2 = (1 + F) \hat{\sigma}_A^2 \) and \( \hat{\sigma}_{DS}^2 = (1 - F^2) \hat{\sigma}_D^2 \). The first expression is also valid when only additive effects are considered and \( \hat{\sigma}_{GS}^2 = (I + F) \hat{\sigma}_A^2 \).

When half-sib and full-sib families are drawn from an inbred population, the families are themselves non-inbreds and the variance among families is expressed by

\[
\hat{\sigma}_F^2 = \theta_1 \hat{\sigma}_A^2 + \theta_2 \hat{\sigma}_D^2 + \theta_1 \hat{\sigma}_{AA}^2 + \theta_2 \hat{\sigma}_{DD}^2 + \theta_1 \theta_2 \hat{\sigma}_{AD}^2 + \theta_1^2 \hat{\sigma}_{AAA}^2 + \cdots
\]

where \( \theta_1 = (1 + F)/4 \) and \( \theta_2 = 0 \) for half-sib families and \( \theta_1 = (1 + F)/2 \) and \( \theta_2 = (1 + F)^2/4 \) for full-sib families, according to adaptation from Cockerham (1963). In Chapter 3 these covariance terms are expressed in terms of covariance between relatives.

Under continuous selfing, assuming only additive effects, the total genetic variance is partitioned into among lines and within lines as follows:

<table>
<thead>
<tr>
<th></th>
<th>( 2F \hat{\sigma}_G^2 )</th>
<th>( (1 - F) \hat{\sigma}_G^2 )</th>
<th>( (1 + F) \hat{\sigma}_G^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where \( \hat{\sigma}_G^2 \) is the total genetic variance in a random mating non-inbred population. Thus when inbreeding is complete (\( F = 1 \)), the genetic variance among lines is twice the genetic variance of the reference population (non-inbred) and no genetic variation is expected within lines. At complete inbreeding (\( F = 1 \)), the additive model is completely valid (since there is no dominance) and is a good approximation for \( F \) slightly less than 1 (highly inbred lines).

### 2.8.7 Variance in a Cross Between Two Populations

The first cross between two distinct populations is not in equilibrium, and its total genetic variance is not linearly related to any of the parent populations. However, total genetic variance can be partitioned into additive and dominance components as follows:

\[
\hat{\sigma}_{A(12)}^2 = pq(a + (s - r)d)^2 + rs(a + (q - p)d)^2 = \frac{1}{2}(\hat{\sigma}_{A12}^2 + \hat{\sigma}_{A21}^2)
\]

\[
\hat{\sigma}_{D(12)}^2 = 4p(1 - p)r(1 - r)d^2 = 4pqrsd^2
\]

according to the notation used in Table 2.6 (Compton et al., 1965).
The two components may be called homologues of additive and dominance variances as defined for one population, i.e., $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. In fact, they equal $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ when $p = r$; i.e., both populations have exactly the same gene frequency. In the above notation we used $\hat{\sigma}_A^{(12)}$ to denote the total additive genetic variance and either $\hat{\sigma}_A^{(12)}$ or $\hat{\sigma}_D^{(12)}$ to denote the subcomponents when either $P_1$ or $P_2$, respectively, is used as the female parent.

When the crossed population is in a half-sib family structure, the variance among families is obtained from Table 2.7 as follows:

$$\hat{\sigma}_{HS}^{(12)} = p^2(r a + s d)^2 + \cdots + q^2(r d - sa)^2 - u_{HS}^{(12)} = (\frac{1}{2})pq[a + (s-r)d]^2$$

and $u_{HS}$ represents the mean of the half-sib population.

If population $P_2$ is used as the female parent

$$\hat{\sigma}_{HS}^{(12)} = (\frac{1}{2})rs[a + (q - p)d]^2 = (\frac{1}{4})\hat{\sigma}_A^{(12)}$$

If both types of families are drawn, their mean variance is $(\frac{1}{4})\hat{\sigma}_A^{(12)} + (\frac{1}{4})\hat{\sigma}_D^{(12)}$; for $p = r$ this equals $(\frac{1}{2})[(\frac{1}{2})\hat{\sigma}_A^{(12)} + (\frac{1}{2})\hat{\sigma}_D^{(12)}] = (\frac{1}{4})\hat{\sigma}_A^2$, which is the variance among half-sib families for one population.

In the same way, genetic variance among full-sib families is obtained from Table 2.8 as follows:

$$\hat{\sigma}_{FS}^{(12)} = p^2r^2a^2 + \cdots + q^2s^2a^2 - u_{FS}^{(12)} = (\frac{1}{2})pq[a + (1 - 2r)d]^2 + pqrsd^2$$

and $u_{FS}$ represents the mean of the full-sib population.

Note that this result will be the same whatever parental population is used as female parent. When $p = r$, then $\hat{\sigma}_{FS}^2 = (\frac{1}{4})\hat{\sigma}_A^2 + (\frac{1}{4})\hat{\sigma}_D^2$ the variance among full-sib families as previously demonstrated for one population.

### 2.9 Means and Variances in Backcross Populations

The same principles apply to backcross populations (see Section 4.12 for more details). Assuming gene frequencies in equilibrium we can describe these populations as follows (two alleles, A and a):

<table>
<thead>
<tr>
<th>PARENTS</th>
<th>AA</th>
<th>×</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>×</td>
<td>Aa</td>
<td>×</td>
</tr>
<tr>
<td>(BC₁)</td>
<td>(½) AA + (½) Aa</td>
<td>(½) Aa + (½) aa</td>
<td>(BC₂)</td>
</tr>
</tbody>
</table>
Then,
\[ X_{BC1} = \left(\frac{1}{2}\right) a + \left(\frac{1}{2}\right) d \]
\[ \hat{\sigma}^2_{BC1} = \left(\frac{1}{2}\right) a^2 + \left(\frac{1}{2}\right) d^2 - \left(\left(\frac{1}{2}\right) a + \left(\frac{1}{2}\right) d\right)^2 \]
\[ = \left(\frac{1}{4}\right) a^2 + \left(\frac{1}{4}\right) d^2 - \left[ \left(\frac{1}{4}\right) a^2 + 2\left(\frac{1}{2}\right)(\frac{1}{2}) ad + \left(\frac{1}{4}\right) d^2 \right] \]
\[ = \left(\frac{1}{4}\right) a^2 + \left(\frac{1}{4}\right) d^2 - \left[ \left(\frac{1}{4}\right) a^2 + \left(\frac{1}{4}\right)d^2 \right] \]
\[ \hat{\sigma}^2_{BC2} = \left(\frac{1}{2}\right) \hat{\sigma}^2_A + \hat{\sigma}^2_D - \left(\frac{1}{2}\right) \text{CovAD} \]
if \( p = q = \frac{1}{2} \)

And then the second backcross population (Note this refers to backcross 1 with a different parent and it is not the result of two backcrosses):
\[ X_{BC2} = \left(\frac{1}{2}\right) d - \left(\frac{1}{2}\right) a \]
\[ \hat{\sigma}^2_{BC2} = \left(\frac{1}{2}\right) d^2 + \left(\frac{1}{2}\right) a^2 - \left(\left(\frac{1}{2}\right) d - \left(\frac{1}{2}\right) a\right)^2 \]
\[ = \left(\frac{1}{4}\right) a^2 + \left(\frac{1}{4}\right) d^2 - \left[ \left(\frac{1}{4}\right) a^2 - 2\left(\frac{1}{2}\right) \left(\frac{1}{2}\right) ad + \left(\frac{1}{4}\right) d^2 \right] \]
\[ = \left(\frac{1}{4}\right) a^2 + \left(\frac{1}{4}\right) d^2 + \left(\frac{1}{2}\right) ad \]
\[ \hat{\sigma}^2_{BC2} = \left(\frac{1}{2}\right) \hat{\sigma}^2_A + \hat{\sigma}^2_D + \left(\frac{1}{2}\right) \text{CovAD} \]
if \( p = q = \frac{1}{2} \)

Therefore, the sum of \( BC_1 + BC_2 = \hat{\sigma}^2_A + 2\hat{\sigma}^2_D \) allows us to eliminate the covariance term. A similar procedure can be done for estimating the variance among and within selfed backcross generations (see Chapter 4).

### 2.10 Heritability, Genetic Gain, and Usefulness Concepts

Heritability is the degree of correspondence between the phenotype and the breeding value of an individual for a particular trait. The best way to determine the breeding value of a plant is to grow and examine its progeny. If the correlation between the breeding value and the phenotype is high then heritability is high (e.g., qualitative, less complex inherited traits). If dominance, epistatic, and environmental effects are large then heritability is low (e.g., quantitative traits). Also, the results of the estimation depend strictly on the population you are working with.

Estimates can be narrow or broad sense depending on which genetic variance is considered. Also, we can determine the heritability on an individual plant basis or on a progeny mean basis depending on the generation used.

Narrow-sense heritability can be defined as the ratio of the additive genetic variance to the phenotypic variance:
\[ \hat{h}^2 = \frac{\hat{\sigma}^2_A}{\hat{\sigma}^2_P} \]
Narrow-sense heritability
Warner (1952) developed a clear estimate of $\hat{\sigma}_A^2$ by applying simple algebra on the generations studied above (for more details see Section 4.12). He showed the following:

\[
\begin{align*}
2\hat{\sigma}_{F_2}^2 &= 2\hat{\sigma}_A^2 + 2\hat{\sigma}_D^2 \\
\hat{\sigma}_{BC_1}^2 + \hat{\sigma}_{BC_2}^2 &= \hat{\sigma}_A^2 + 2\hat{\sigma}_D^2 \\
&= \hat{\sigma}_A^2
\end{align*}
\]

Therefore,

\[
\hat{h}^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_p^2}
\]

is a narrow-sense heritability estimate on an individual plant basis.

In maize, heritability estimates for grain yield vary from less than 0.1 (10%) when it is based on individual plants grown in one location (e.g., mass selection) to >0.8 (80%) when it is based on inbred progenies grown across locations with full-sibs and half-sibs having intermediate values.

There are several methods to estimate heritability on an individual plant basis. Different estimates can be obtained using the same populations:

(a) Burton (1951)

\[
\hat{h}^2 = \frac{\hat{\sigma}_{F_2}^2 - \hat{\sigma}_{F_1}^2}{\hat{\sigma}_{F_2}^2}
\]

broad sense

(b) Warner (1952)

\[
\hat{h}^2 = \frac{[2\hat{\sigma}_{F_2}^2 - (\hat{\sigma}_{BC_1}^2 + \hat{\sigma}_{BC_2}^2)]/\hat{\sigma}_{F_2}^2}{\hat{\sigma}_{F_2}^2}
\]
narrow sense

(c) Mahmud and Kramer (1951)

\[
\hat{h}^2 = \frac{[\hat{\sigma}_{F_2}^2 - (\hat{\sigma}_{P_1}^2 \times \hat{\sigma}_{F_2}^2)]/\hat{\sigma}_{F_2}^2}{\hat{\sigma}_{F_2}^2}
\]
broad sense

(d) Weber and Moorthy (1952)

\[
\hat{h}^2 = \left[\frac{\hat{\sigma}_{F_2}^2 - \left(\hat{\sigma}_{P_1}^2 \times \hat{\sigma}_{F_2}^2 \times \hat{\sigma}_{F_1}^2\right)^{1/3}}{\hat{\sigma}_{F_2}^2}\right] / \hat{\sigma}_{F_2}^2
\]
broad sense
Going back to our generations:

\[ \hat{h}^2_\text{F2} = \frac{\hat{\sigma}^2_\text{F2} - \hat{\sigma}^2_\text{W}}{\hat{\sigma}^2_\text{F2}} \]

Usually less than 0.1
Heritability in broad sense/individual plant basis
It is common to use only one location
\( \hat{\sigma}^2_\text{W} \) environmental effects (can be estimated)

\[ \hat{h}^2_\text{F3} = \frac{\hat{\sigma}^2_\text{F3}}{\hat{\sigma}^2_e/r + \hat{\sigma}^2_\text{F3}} \]
Can be 0.75
Heritability in broad sense
It is common to use several observations per line
Based on S1 progeny means

\[ \hat{h}^2_\text{F4} = \frac{\hat{\sigma}^2_\text{F4}}{\hat{\sigma}^2_e/r + \hat{\sigma}^2_\text{F4}} \]
Can be 0.87 (50% more additive variance)
Heritability in broad sense
It is common to use several observations per line
Based on S2 progeny means

If progenies (g) are evaluated in replicated (r) trials in different environments (e), heritability in the broad sense based on progeny means is:

\[ \hat{h}^2 = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2_e/r + \hat{\sigma}^2_g/e + \hat{\sigma}^2_ge} \]
Heritability in broad sense
Progeny mean basis

As expected heritability contributes to \textit{genetic gain} (\( \Delta G \)) as follows:

\[ \Delta G = \hat{h}^2S \]
and \( S = \text{selection differential} = (\bar{X}_s - \bar{X}) \)

being \( \bar{X}_s \) Mean of progeny selected and \( \bar{X} \) Mean of the overall population

Table 2.14 shows an example of an intra-population recurrent selection program on NDSAB genetically broad-based population improved by 12 cycles of modified ear-to-row selection plus two cycles of full-sib recurrent selection. A heritability index was utilized for selecting the top progenies based on grain yield (YIELD), grain moisture at harvest (MOIST), test weight (TWT), and stalk lodging resistance (SL). TRT refers to treatment number and PI to a performance index including grain yield and grain moisture at harvest.

Breeders want to identify populations with high mean and large genetic variance. A function was created to combine information on the mean performance and genetic variance of a population and is called \textit{usefulness criterion} (U):

\[ U = \bar{X} + \Delta G \]
Table 2.14 Recurrent selection program including 200 full-sib progenies grown across three ND locations arranged in an augmented unreplicated experimental design. Only the top 16 full-sib progenies that were included to form cycle 14 are included.

<table>
<thead>
<tr>
<th>Pedigree</th>
<th>TRT</th>
<th>Yield</th>
<th>PI</th>
<th>MSTR</th>
<th>SL</th>
<th>TWT</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-142</td>
<td>145</td>
<td>81.3</td>
<td>118.5</td>
<td>24.5</td>
<td>4.5</td>
<td>53.5</td>
<td>7.42</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-3</td>
<td>12</td>
<td>70.5</td>
<td>120.2</td>
<td>21.3</td>
<td>2.7</td>
<td>53.6</td>
<td>6.85</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-106</td>
<td>54</td>
<td>74.0</td>
<td>125.3</td>
<td>21.0</td>
<td>12.5</td>
<td>56.3</td>
<td>4.73</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-76</td>
<td>95</td>
<td>71.9</td>
<td>109.4</td>
<td>23.6</td>
<td>6.9</td>
<td>52.4</td>
<td>4.54</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-112</td>
<td>43</td>
<td>66.5</td>
<td>107.2</td>
<td>22.1</td>
<td>6.2</td>
<td>55.4</td>
<td>4.06</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-73</td>
<td>80</td>
<td>63.9</td>
<td>113.8</td>
<td>20.0</td>
<td>8.1</td>
<td>56.8</td>
<td>3.99</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-119</td>
<td>55</td>
<td>64.3</td>
<td>102.5</td>
<td>22.3</td>
<td>4.3</td>
<td>57.1</td>
<td>3.97</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-4</td>
<td>3</td>
<td>67.2</td>
<td>108.2</td>
<td>21.8</td>
<td>9.1</td>
<td>55.1</td>
<td>3.52</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-17</td>
<td>18</td>
<td>68.0</td>
<td>114.9</td>
<td>20.8</td>
<td>11.5</td>
<td>57.7</td>
<td>3.52</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-24</td>
<td>129</td>
<td>74.0</td>
<td>121.7</td>
<td>21.7</td>
<td>15.1</td>
<td>52.8</td>
<td>3.50</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-14</td>
<td>5</td>
<td>55.4</td>
<td>100.8</td>
<td>19.7</td>
<td>2.9</td>
<td>58.1</td>
<td>3.47</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-97</td>
<td>64</td>
<td>67.4</td>
<td>99.7</td>
<td>24.0</td>
<td>5.6</td>
<td>52.6</td>
<td>3.46</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-61</td>
<td>84</td>
<td>59.0</td>
<td>98.9</td>
<td>21.6</td>
<td>3.1</td>
<td>56.7</td>
<td>3.35</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-36</td>
<td>122</td>
<td>79.2</td>
<td>112.1</td>
<td>25.3</td>
<td>13.8</td>
<td>51.1</td>
<td>3.32</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-94</td>
<td>58</td>
<td>68.4</td>
<td>107.0</td>
<td>23.0</td>
<td>9.3</td>
<td>54.5</td>
<td>3.07</td>
</tr>
<tr>
<td>NDSAB(MER-FS)C13 SYN1-92</td>
<td>70</td>
<td>61.7</td>
<td>94.3</td>
<td>23.4</td>
<td>3.0</td>
<td>53.4</td>
<td>3.04</td>
</tr>
</tbody>
</table>

\[ \bar{X} = \{(0.30 \times \text{Yield}) + (0.69 \times \text{TWT}) - (0.58 \times \text{MSTR}) - (0.33 \times \text{SL})\} \]

Index = \[(0.30 \times \text{Yield}) + (0.69 \times \text{TWT}) - (0.58 \times \text{MSTR}) - (0.33 \times \text{SL})\]

2.11 Generation Mean Analysis

We will now introduce our first genetic analysis with development of progenies. This is a type of genetic analysis that can be useful for preliminary studies. Several models have been developed for the analysis of generation means. It is important to note that these models estimate the relative importance of genetic effects based upon the means of different generations and not based on their variances. Genetic variances are determined from the summation of squared effects for each locus. This type of genetic analysis does not involve development of progenies that have a family structure of sib-ships as in Chapter 4, but it includes genetic populations (or generations) that are similar to those characterized by the special case of \( p = q = 0.5 \) (e.g., \( F_2 \) populations). Instead of estimating genetic variation within generations, we will concern ourselves with relative genetic effects estimated from the means of different generations. Mather (1949) presented several generation comparisons to test for additiveness of genetic effects for estimation of \( \sigma_A^2 \) and \( \sigma_D^2 \). If the scale of measurement deviated from additivity, he suggested a transformation to make the effects additive. The generation models were extended to include estimation of epistatic effects. Several models have been developed for analysis of generation means (Anderson and Kempthorne, 1954; Hayman, 1958, 1960; Van der Veen, 1959; Gardner and Eberhart, 1966). Because of similarity in nomenclature we
will use the Hayman (1958, 1960) model to illustrate the type of genetic information obtained from generation mean analyses. As an example, consider the generations produced from the cross of two inbred lines (e.g., special case of $p = q = 0.5$). For estimation of genetic effects we will use the means of each generation rather than develop progenies within the segregating generations.

Hayman (1958) defined his base population as the F$_2$ population resulting from a cross of two inbred lines. If they differ by any number of unlinked loci, expectations of parents and their descendant generations in terms of genetic effects relative to the F$_2$ generation are as follows:

$$
\begin{align*}
P_1 &= m + a - (\gamma_2)d + aa - ad + (\gamma_4)dd \\
P_1 &= m - a - (\gamma_2)d + aa + ad + (\gamma_4)dd \\
F_1 &= m + (\gamma_2)d + (\gamma_4)dd \\
F_2 &= m \\
F_3 &= m - (\gamma_6)d + (\gamma_{16})dd \\
F_4 &= m - (\gamma_6)d + (\gamma_{64})dd \\
F_5 &= m - (\gamma_{16})d + (\gamma_{256})dd \\
F_6 &= m - (\gamma_{32})d + (\gamma_{1024})dd \\
BC_1 &= m + (\gamma_2)a + (\gamma_4)aa \\
BC_2 &= m - (\gamma_2)a + (\gamma_4)aa \\
BC_2^1 &= m + (\gamma_4)a + (\gamma_{16})aa \\
BC_2^2 &= m - (\gamma_4)a + (\gamma_{16})aa \\
BS_1 &= m + (\gamma_2)a - (\gamma_4)d + (\gamma_4)aa - (\gamma_4)ad + (\gamma_{16})dd \\
BS_2 &= m - (\gamma_2)a - (\gamma_4)d + (\gamma_4)aa - (\gamma_4)ad + (\gamma_{16})dd
\end{align*}
$$

or in general the observed mean $= m + \alpha a + \beta d + \alpha^2 aa + 2\alpha \beta ad + \beta^2 dd$, where $\alpha$ and $\beta$ are the coefficients of $a$ and $d$. Because the F$_2$ mean is $(\gamma_2)d$ and the F$_1$ mean is equal to $d$, the F$_1$ mean relative to the F$_2$ mean has added an increment of $(\gamma_2)d$. Terms $a$ and $d$ are the same as those illustrated in the special case of $p = q = 0.5$, where $a$ indicates additive effects and $d$ indicates dominance effects. Hayman (1958) used lowercase letters to indicate summation over all loci by which the two inbred lines differ. Thus $a$ measures the pooled additive effects; $d$ the pooled dominance effects; and $aa$, $ad$, and $dd$ the pooled digenic epistatic effects.

The different generations listed can be produced rather easily for cross- and self-pollinated species. Hand pollinations are necessary for all generations in maize, but selfing generations can be obtained naturally in self-pollinated species. Bulks of progenies of each generation are evaluated in replicated experiments repeated over environments, and generation means can be determined for traits under study. In growing different generations, one should be cognizant of two important considerations in order to have valid estimates of the generation means:

1. Sufficient sampling of segregating generations is necessary to have a representative sample of genotypes. In parental and F$_1$ generations no sampling is involved, but F$_2$, F$_3$, F$_4$, . . . , and backcross generations will be segregating and sample size has to be considered.
In maize it is necessary to consider the level of inbreeding of each generation, and it becomes necessary to have sufficient border rows in experimental plots to minimize competition effects of adjacent plots.

Several different possibilities exist for the type and number of generations that can be included in a generation mean experiment. If the two parents and the \( F_1, F_2, \) and \( F_3 \) generations are evaluated, we have five means for comparison. Expectations of each generation can be determined and a least-squares analysis made to estimate \( m, a, \) and \( d \) with a fair degree of precision. For this simple experiment we can also make a goodness-of-fit test (observed means compared with predicted means) to determine the sufficiency of the model for \( m, a, \) and \( d \) to explain the differences among the generation means.

Letting \( m = \) general mean, \( a = \) sum of signed additive effects, and \( d = \) signed dominance effects, we have the following expressions with \( F_2 \) as the base population:

\[
\begin{align*}
P_1 &= m + a \\
F_1 &= m + \left( \frac{1}{2} \right) d \\
F_2 &= m \\
F_3 &= m - \left( \frac{1}{4} \right) d \\
\end{align*}
\]

By use of the technique suggested by Mather (1949), the five equations can be reduced to the following normal equations:

\[
\begin{align*}
5m + \left( \frac{1}{4} \right) d &= Q_1 (P_1 + P_2 + F_1 + F_2 + F_3) \\
2a &= Q_2 (P_1 - P_2) \\
\left( \frac{1}{4} \right) m + \left( \frac{5}{16} \right) d &= Q_3 (F_1 + F_2)
\end{align*}
\]

In matrix form the set of equations is equal to

\[
\begin{bmatrix}
5 & 0 & \frac{1}{4} \\
0 & 2 & 0 \\
\frac{1}{4} & 0 & \frac{5}{16}
\end{bmatrix}
\begin{bmatrix}
m \\
a \\
d
\end{bmatrix}
= \begin{bmatrix}
Q_1 \\
Q_2 \\
Q_3
\end{bmatrix}
\]

Solving for parameters \( m, a, \) and \( d, \) we get the following estimates:

\[
\begin{align*}
\hat{m} &= \left( \frac{5}{24} \right) Q_1 - \left( \frac{1}{6} \right) Q_3 \\
\hat{a} &= \left( \frac{1}{2} \right) Q_2 \\
\hat{d} &= -\left( \frac{1}{6} \right) Q_1 + \left( \frac{10}{3} \right) Q_3
\end{align*}
\]

The estimates of \( m, a, \) and \( d \) can be inserted in the predicted values and can be compared with the observed values for each generation. If the squares of deviations of the expected from the observed are significant, the three estimated parameters
are not sufficient to explain differences among the generation means; this is a
goodness-of-fit test for epistasis and/or linkage to determine if the three parameters
included are sufficient or if more are needed. The best procedure would be to
sequentially fit the successive models starting with the mean and add one term with
each successive fit. Tests of residual mean squares can be made for each model to
determine how much of the total variation among generations is explained by dif-
ferent parameters in the model. High-speed computers facilitate such computations,
and a weighed least-squares analysis can be done rather easily.

The similarity of genetic populations included for this simple generation mean
experiment and genetic populations used for estimation of genetic variances for
the special case of \( p = q = 0.5 \) is obvious. If necessary measurements have been
made for the different genetic populations, we can also obtain the following sets of
equations:

\[
\begin{align*}
\text{Variance among F}_2 \text{ individuals} & = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + E_1 \\
\text{Variance among F}_3 \text{ progeny means} & = \hat{\sigma}_A^2 + \langle \iota_1 \rangle \hat{\sigma}_D^2 + E_2 \\
\text{Variance within F}_3 \text{ progenies} & = \langle \iota_2 \rangle \hat{\sigma}_A^2 + \langle \iota_2 \rangle \hat{\sigma}_D^2 + E_1 \\
\text{Covariance between F}_2 \text{ individuals and F}_3 \text{ progeny means} & = \hat{\sigma}_A^2 + \langle \iota_2 \rangle \hat{\sigma}_D^2 \\
\text{Variance among parents and F}_1 \text{ individuals} & = E_1 \\
\text{Experimental error} & = E_2
\end{align*}
\]

Six equations are available for estimation of two heritable and two non-heritable
sources of variation. Direct estimates of \( E_1 \) and \( E_2 \) are available, but an unweighted
(or preferably a weighted) least-squares analysis can be used to estimate the four
parameters (\( \hat{\sigma}_A^2, \hat{\sigma}_D^2, E_1, \) and \( E_2 \)) from the six equations. If one estimates the genetic
effects \( a \) and \( d \) (by use of generation mean analysis) and the genetic variances \( \hat{\sigma}_A^2 \)
and \( \hat{\sigma}_D^2 \), there probably will be little relation in the magnitude of the two sets of
estimates. This should be expected because in the first instance we are estimating the
sum of the signed genetic effects, whereas in the second instance we are estimating
variances that are the squares of the genetic effects. For maize it seems that estimates
of the \( d \) effects are usually greater, especially if the F\(_1\) generation is included. On
the other hand, estimation of genetic variances in maize usually shows that estimates
of \( \hat{\sigma}_A^2 \) are similar to or greater than estimates of \( \hat{\sigma}_D^2 \). The expression of heterosis in
F\(_1\) crosses of two inbred lines of maize probably has a much greater effect on the
estimate of \( d \) for maize than for many other crop species.

Limitations of the generation mean analysis if the model includes epistatic effects
were discussed by Hayman (1960). Briefly, if the residuals are not significantly dif-
f erent from zero after \( m, a, \) and \( d \) are fitted, we have unique estimates for \( a \) and
\( d \). However, if it is necessary to include epistatic effects in the model, estimates
digenic epistatic effects are unique but estimates of \( a \) and \( d \) are confounded
with some of the epistatic effects. Hence if epistatic effects are not present, esti-
mates of \( a \) and \( d \) effects are meaningful and unbiased by linkage disequilibrium;
if epistatic effects are present, estimates of \( a \) and \( d \) effects are biased by epistatic
effects and linkage disequilibrium (if present). Estimation of digenic epistatic effects
is unbiased if linkage of interacting loci and higher order epistatic effects are absent. Because of the bias in the estimates of \( a \) and \( d \) effects, when a model that includes epistatic effects is used, the relative importance of \( a \) and \( d \) effects vs. epistatic effects cannot be directly assessed. Some indication of their relative importance may be gained by comparing residual sums of squares after fitting the three-parameter \((m, a, d)\) and the six-parameter \((m, a, d, aa, ad, dd)\) models.

It seems that the primary function of generation mean analysis is to obtain some specific information about a specific pair of lines. How useful the information obtained from generation mean analysis is to the maize breeder is not obvious. For quantitative traits the estimates of genetic effects would be quite different for different pairs of lines, depending on the relative frequency of opposing and reinforcing effects for the specific pair of lines studied. The cancellation of opposing effects may confound interpretations, but a complete diallel of interested lines could be used to determine which have opposing and reinforcing effects. Generation mean analysis is amenable for use in self-pollinated species because limited hand pollinations are required to produce the different generations; hence generation mean analysis may provide some information on the relative importance of non-additive genetic effects for the justification of a hybrid breeding program. The relative importance of dominance effects could be determined by comparing different generations derived from the \( F_1 \) generation, which would involve only the cross of the two parents to produce the \( F_1 \) generation. For maize, however, controlled pollinations would be necessary for all generations.

Generation mean analysis has some advantages and disadvantages in comparison with mating designs used for estimation of genetic components of variance (see Chapter 4). Because we are working with means (first-order statistics) rather than variances (second-order statistics), the errors are inherently smaller. We can rather easily extend generation mean analysis to more complex models that include epistasis, but the main effects \((a \text{ and } d)\) are not unique when epistatic effects are present. Generation mean analysis is equally applicable to cross- and self-pollinating species. Smaller experiments are required for generation mean analysis to obtain the same degree of precision. However, an estimate of heritability cannot be obtained and one cannot predict genetic advance because estimates of genetic variances are not available. Cancellation of effects may be a significant disadvantage because, say, dominance effects may be present but opposing at various loci in the two parents and cancel each other. Generation mean analysis does not reveal opposing effects, but this may be overcome to some extent by a balanced set of diallel crosses.

This discussion for generation mean analysis is restricted to the use of parents being either inbred or pure lines, i.e., relatively homozygous and homogeneous. Generation mean analysis has been extended to populations generated from parents that are not homozygous. Robinson and Cockerham (1961) presented an analysis for two varieties and \( n \) alleles at each locus. Gardner and Eberhart (1966) included \( n \) varieties with only two alleles at a locus. The analysis by Robinson and Cockerham (1961) is orthogonal in the partitioning of sums of squares, and tests can be made to determine the presence of non-additive effects. All the generation mean
analyses provide information on the relative importance of genetic effects, but the information in most instances may not be useful to applied breeders, particularly those that are conducting long-term selection programs.

In summary, if we consider the generations produced from the cross of two inbred lines, Hayman (1958) defines the F2 generation as his base or reference population. Therefore, if inbred lines differ by any number of unlinked loci, expectations of parents and their descendant generations in terms of genetic effects relative to the F2 generation are based on the following model:

\[ Y_{ijkl} \text{(observed mean)} = m + \alpha_i + \beta_j + \alpha_k^2 + 2\alpha\beta_{ij} + \beta_1^2 \quad \text{or} \]

\[ Y = m + \alpha_a + \beta_d + \gamma_{aa}^2 + 2\alpha\beta_{ad} + \beta_{dd}^2 \]

following notation of Gamble (1962a, b)

‘\(a\)’ indicates pooled additive effects across loci
‘\(d\)’ indicates pooled dominance effects across loci
‘\(aa\),’ ‘\(ad\),’ and ‘\(dd\)’ indicate pooled digenic epistatic effects

Considering the means for each generation as follows:

<table>
<thead>
<tr>
<th>Generation</th>
<th>Mean</th>
<th>Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>(a)</td>
<td>(m + a - (\frac{1}{2})d + aa - ad + (\frac{1}{4})dd)</td>
</tr>
<tr>
<td>P2</td>
<td>(-a)</td>
<td>(m - a - (\frac{1}{2})d + aa + ad + (\frac{1}{4})dd)</td>
</tr>
<tr>
<td>F1</td>
<td>(d)</td>
<td>(m + (\frac{1}{2})d + (\frac{1}{4})dd)</td>
</tr>
<tr>
<td>F2</td>
<td>((\frac{1}{2})d)</td>
<td>(m)</td>
</tr>
<tr>
<td>BC1</td>
<td>((\frac{1}{2})a + (\frac{1}{2})d)</td>
<td>(m + (\frac{1}{2})a + (\frac{1}{4})aa)</td>
</tr>
<tr>
<td>BC2</td>
<td>(-((\frac{1}{2})a + (\frac{1}{2})d))</td>
<td>(m - (\frac{1}{2})a + (\frac{1}{4})aa)</td>
</tr>
</tbody>
</table>

More generations can be included. The more generations we include the better the estimates (e.g., less error, more precision) but more experiments might be needed.

Therefore, we can estimate gene effects when epistasis is present with information of generation means:

\[ \text{‘} a \text{’} = \overline{BC}_1 - \overline{BC}_2 \]
\[ \text{‘} d \text{’} = \overline{F}_1 - 4\overline{F}_2 + 2\overline{BC}_1 + 2\overline{BC}_2 - (\frac{1}{2})\overline{P}_1 - \overline{P}_2 \]
\[ \text{‘} aa \text{’} = 2\overline{BC}_1 + 2\overline{BC}_2 - 4\overline{F}_2 \]
\[ \text{‘} ad \text{’} = (\frac{1}{2})\overline{P}_2 - (\frac{1}{2})\overline{P}_1 + \overline{BC}_1 - \overline{BC}_2 \]
\[ \text{‘} dd \text{’} = \overline{P}_1 + \overline{P}_2 + 2\overline{F}_1 + 4\overline{F}_2 - 4\overline{BC}_1 - 4\overline{BC}_2 \]

In this case we do not expect deviations since the number of unknown estimates equals the number of generations used in the model.
2.11.1 Assumptions for Analysis

1) Two alleles per locus.
2) Most positive alleles in P1 and most negative alleles in P2.
3) No linkage of interacting loci.
4) No trigenic or higher order epistasis.
5) F2 is the reference population.
6) Sufficient sampling of segregating generations (representative sample of genotypes).
7) No competition effects among generations at different levels of inbreeding.

2.11.2 Limitations of This Analysis

1) Unlike variances, means can estimate neither heritability nor genetic gain for prediction.
2) Genetic effects are always summing or subtracting. Therefore, there is a cancellation of effects that are not detected.

2.11.3 Advantages of This Analysis

1) Useful for preliminary studies, i.e., is there enough dominance to have good hybrids? Allows to study a new or unknown population.
2) Means are estimated with greater precision than variances.
3) Can be extended to more complex models.

References

References


Chapter 3
Resemblance Between Relatives

3.1 Introduction

The degree of relationship between relatives depends on genetic resemblance. They are very important to breeders since they are not only used to estimate the genetic variances of reference plant populations but also used in breeding methods used for selection. All breeding methods deal to some extent with resemblance between relatives.

Progress from selection (recurrent or not) is directly proportional to the degree of resemblance between progenies and the selected parent. Besides parent and offspring, other types of relationships are useful in several aspects of quantitative genetics and breeding procedures. Therefore covariance between relatives is important in modern plant breeding for at least two reasons: (1) In most instances covariances between relatives can be expressed in terms of components of genetic variance of the reference population. On the other hand, the variance among families can in some instances be expressed as linear functions of covariance between relatives, thus allowing the estimation of components of genetic variance by using appropriate experimental and breeding designs. (2) The expected progress from selection depends basically on the degree of relationship (i.e., covariance) between the unit of selection (individuals or families) and the individuals, descendant from the selected parents.

The first reports of covariation and correlation between relatives were given by Fisher (1918) and Wright (1921), although previous studies were reported in the early part of the 20th century (Pearson, 1904; Yule, 1906; Weinberg, 1908, 1910) as mentioned by Kempthorne (1954). A general theory for the genetic interpretation of covariance between relatives was given by Cockerham (1954) and Kempthorne (1954, 1955) that included epistasis in addition to additive and dominance effects. Both authors used a factorial approach to the partition of genetic variability, assuming no linkage, and Kempthorne’s (1954) model is not restricted to the number of alleles per locus. Linkage effects were further presented by Cockerham (1956) and Schnell (1963). Therefore, Fisher and Wright gave the first reviews on the correlation between relatives while Cockerham and Kempthorne gave the most common cases of the covariance of relatives and their genetic interpretation.
A summary of the topic is presented across authors such as those of Fisher (1918), Wright (1921), Cockerham (1954, 1963), Kempthorne (1954, 1957), Falconer and Mackay (1996).

### 3.2 Theoretical Basis of Covariance

Basically, the covariance between relatives depends on the genetic resemblance between them. Environmental effects are usually expected to be uncorrelated between relatives so that covariance deals with only genotypic values. General theories are limited to some reference population, which is assumed to be in Hardy–Weinberg equilibrium, and the results are extended to any number of loci and any number of alleles per locus. Effects of inbreeding may be included for either the relatives or the reference population or both.

As a simple example consider the covariance between parent and offspring:

<table>
<thead>
<tr>
<th>Parents</th>
<th>Offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>$X_2$</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>$X_i$</td>
<td>$Y_i$</td>
</tr>
<tr>
<td>$X_n$</td>
<td>$X_n$</td>
</tr>
</tbody>
</table>

First, let us find the causation of resemblance between $X_i$ and $Y_i$. The $X_i$ genotype may be expressed by $G_X = A_X + D_X$, where $G_X$ is measured as the deviation from the population mean, $A_X$ is the breeding value or additive genotypic value of $X_i$, and $D_X$ is the dominance deviation from the population mean. Since $A$ and $D$ are taken as deviations from the population mean, specification of a reference population is an obvious requirement.

Consider an offspring $Y_i$, which represents the expected value or average value of the progeny. It has been seen that the average value of a progeny is one-half the breeding value of the parent, i.e., $G_Y = \left(\frac{1}{2}\right)A_X$. A dominance effect is not assigned in $Y$ genotypic value because it has no relation with dominance effect in $X$. In other words, dominance effects are not transferable through gametes (haploid entities) but are recreated at random in the offspring. Hence the parent–offspring covariance is that of individual genotypes with one-half of their respective breeding values, i.e., $\text{Cov}(X, Y) = E\{(A_X + D_X)\left(\frac{1}{2}\right)\} = \left(\frac{1}{2}\right)E(A_X^2) = \left(\frac{1}{2}\right)\sigma^2_A$ because the expected value $E$ of the squared breeding value is the additive genetic variance. Because $A$ and $D$ are uncorrelated, the expected value of the term $AD = 0$.

The above calculations of covariance between relatives are known as the direct method (Kempthorne, 1957). However, the resemblance between relatives derives from the probability that two relatives have identical alleles (one or two) at a given
locus. By identical we mean identical by descent. An identical allele present in two relatives is an exact copy of an allele present in a common background.

In the case of the parent–offspring relationship, the probability is that an allele in the offspring is identical by descent to the same allele in the parent. In this case the probability that any individual in the offspring has, at a given locus, both alleles identical to those in the parent is zero, since one of these alleles comes from another parent taken at random from the population. This is the reason why a dominance effect in the offspring cannot be identical to that in the parent.

It can also be seen that the probability of both relatives having the same alleles (identical by descent) at a given locus is different from zero only when they have two or more common ancestors. For example, the covariance between full-sibs is \((\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D\) for a non-inbred reference population. The dominance component arises because two full-sibs taken at random from a progeny may have both alleles common at a given locus with a probability of \(\frac{1}{4}\). However, the possible common dominance effects in both full-sibs are not related to the dominance effects in the parent as far as identity by descent is concerned. For example, consider two parents with genotypes \(A_iA_j\) and \(A_kA_1\). The possible genotypes in the full-sib family are as follows:

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_iA_k)</td>
<td>(\frac{1}{4})</td>
</tr>
<tr>
<td>(A_iA_1)</td>
<td>(\frac{1}{4})</td>
</tr>
<tr>
<td>(A_jA_k)</td>
<td>(\frac{1}{4})</td>
</tr>
<tr>
<td>(A_jA_1)</td>
<td>(\frac{1}{4})</td>
</tr>
</tbody>
</table>

If a pair of individuals is taken at random, the probability that both have the same alleles at locus A is \(P(A_iA_k, A_iA_k) + P(A_iA_1, A_iA_1) + P(A_jA_k, A_jA_k) + P(A_jA_1, A_jA_1) = \frac{1}{4}\). When the full-sibs are identical twins, the probability that they have the same alleles at a locus is 1 and their covariance is \(\hat{\sigma}_A^2 + \hat{\sigma}_D^2\).

There is another case where the dominance effects in the offspring are related to that of parents. It occurs in asexually propagated species where progeny genotypes are identical to their parents and identical among themselves. In self-fertilizing homozygous species, progeny genotypes are also identical to their parents, but in this case there is no dominance because there are no heterozygotes.

### 3.3 Covariance Between Relatives as a Linear Function of Genetic Variances

The most common cases of covariance between relatives are based mainly on the references of Cockerham (1954, 1963) and Kempthorne (1954, 1957). In general, these cases may be classified as follows:
Non-inbred relatives from either a non-inbred or an inbred reference population.

Inbred relatives when either gene frequency is one-half or arbitrary.

3.3.1 Non-inbred Relatives

3.3.1.1 Relatives from a Non-inbred Reference Population

A very common situation in maize breeding involves non-inbred relatives from a non-inbred population. A general method for finding covariance between relatives when epistasis is present is given by Kempthorne (1954). According to his method the covariances depend on two coefficients, \( \Phi \) and \( \Phi' \), which vary according to the genetic resemblance between relatives (Malecot, 1969). The coefficient \( \Phi \) is defined as the probability of two relatives receiving by descent the same genes from one or more common ancestors tracing back through one parent; \( \Phi' \) is defined in the same way for the parent of the opposite sex. The coefficients for the components of variance are found, using \( \hat{\sigma}^2_x \) as a general notation, where \( x \) contains \( r \) A subscripts (for additive effects) and \( s \) D subscripts (for dominance effects). Using the following formula for the covariance of relatives, \( \tilde{\text{Cov}}(X, Y) = \sum [(\varphi + \varphi')/2]^r \varphi \varphi'^s \), the coefficients \( C \) of \( \hat{\sigma}^2_A \) and \( \hat{\sigma}^2_D \) can be obtained.

For example, in the covariance between full-sibs, \( \Phi = \frac{1}{2} \) and \( \Phi' = \frac{1}{2} \), so

\[
\begin{align*}
\hat{\sigma}^2_A &\quad r = 1, s = 0 \\
\hat{\sigma}^2_D &\quad r = 0, s = 1 \\
\hat{\sigma}^2_{AA} &\quad r = 2, s = 0 \\
\hat{\sigma}^2_{DD} &\quad r = 0, s = 2 \\
\hat{\sigma}^2_{AD} &\quad r = 1, s = 1
\end{align*}
\]

The expectations of the genetic components of variance for the covariance of full-sibs become \( (\frac{1}{2}) \hat{\sigma}^2_A + (\frac{1}{4}) \hat{\sigma}^2_D + (\frac{1}{4}) \hat{\sigma}^2_{AA} + (\frac{1}{8}) \hat{\sigma}^2_{AD} + (\frac{1}{16}) \hat{\sigma}^2_{DD} \).

Therefore, we can summarize that according to Kempthorne the covariance of relatives depends on two coefficients:

\[
\Phi = \text{Probability of two relatives receiving by descent the same alleles from one or more common ancestors tracing back through one parent.}
\]

\[
\Phi' = \text{Probability of two relatives receiving by descent the same alleles from one or more common ancestors tracing back through the second parent.}
\]

The coefficient for the components of variance can be obtained through the following formula:

\[
\tilde{\text{Cov}} = \frac{\Phi + \Phi'}{2} \hat{\sigma}^2_A + \Phi \times \Phi' \hat{\sigma}^2_D
\]

\[
\hat{\sigma}^2_x = \frac{C}{\Phi} + \frac{C}{\Phi'}
\]

\[
\Phi' = \frac{1}{2} \hat{\sigma}^2_A + \frac{1}{4} \hat{\sigma}^2_D + \frac{1}{8} \hat{\sigma}^2_{AA} + \frac{1}{16} \hat{\sigma}^2_{DD}
\]
As an example, below is a graphic demonstration of the covariance between full-sibs. Full-sibs are progenies developed with common parents.

\[
\begin{array}{c}
\Phi = P(e = g) \\
= P(a = e; a = g) + P(b = e; b = g) \\
= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \\
\Phi = \frac{1}{2} \\
\Phi' = P(f = h) \\
= P(c = f; c = h) + P(d = f; d = h) \\
= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \\
\Phi = \frac{1}{2}
\end{array}
\]

Therefore, by definition the covariance between full-sibs is

\[
\widehat{\text{Cov}_{FS}} = (\frac{1}{2}) \hat{\sigma}_A^2 + (\frac{1}{4}) \hat{\sigma}_D^2 \quad \text{if} \quad F = 0
\]

These are the expectations of the genetic components of variance for the covariance of full-sibs from non-inbred relatives derived from a non-inbred reference population.

Full-sib progenies will always be non-inbred regardless of \(F\) but \(F\) influences parents (e.g., if they are partially or fully inbred) as explained later.

Another example shown below is a graphic demonstration of the covariance between half-sibs.

Half-sibs are progenies developed with one parent in common.
What is the probability of \( c = e \) (identical by descent)?

\[
\Phi = P(c = e) = P(a = c; a = e) + P(b = c; b = e) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}
\]

\( \Phi' = 0 \)  
Males are different. Alleles identical by descent are not possible.

Therefore, by definition the covariance between half-sibs is

\[
\hat{\text{Cov}}_{\text{HS}} = (\frac{1}{4}) \hat{\sigma}_A^2 \quad \text{if } F = 0
\]

### 3.3.1.2 Relatives from an Inbred Reference Population

For the examples of the covariances of full-sibs and half-sibs the method is straightforward.

What will happen if relatives come from an inbred reference population?

Cockerham developed a general formula that includes the degree of inbreeding \( (F) \) in the reference population. For the case of full-sibs the following formula applies:

\[
\hat{\text{Cov}}_{\text{FS}} = [(1 + F)/2] \hat{\sigma}_A^2 + [(1 + F)/2]^2 \hat{\sigma}_D^2
\]

Therefore,

\[
\hat{\text{Cov}}_{\text{FS}} = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 \quad \text{if } F = 1
\]

In the case of half-sibs

\[
\hat{\text{Cov}}_{\text{FS}} = [(1 + F)/4] \hat{\sigma}_A^2 \quad \text{or} \quad (\frac{1}{2}) \hat{\sigma}_A^2 \quad \text{if } F = 1
\]

Similar procedures are applied for different types of covariance of relatives (e.g., parent–offspring and/or mid-parent–offspring regressions). In these cases, prolific plants (more than one ear per plant) can be used to develop parents and offsprings during the same season.
\[ \hat{\text{Cov}}_{\text{PO}} = \left[ (1 + F)/2 \right] \hat{\sigma}^2_\delta \text{ or } \hat{\sigma}^2_\delta \text{ if } F = 1 \]

When an inbred reference population is under consideration, Cockerham’s (1963) general formula can be used:

\[ \hat{\text{Cov}}(x, y) = \theta_1 \hat{\sigma}^2_\delta + \theta_2 \hat{\sigma}^2_D + \theta_1 \hat{\sigma}^2_{\delta A} + \theta_2 \hat{\sigma}^2_{AD} + \theta_1 \theta_2 \hat{\sigma}^2_{AA} + \cdots \]

where \( \theta_1 \) and \( \theta_2 \) are expressed as a function of \( F \). Obviously, it also includes the case of a non-inbred reference population (\( F = 0 \)). Both Cockerham’s (1963) and Kempthorne’s (1954) coefficients are presented in Table 3.1 for the most common covariances of relatives; \( F \) is the coefficient of inbreeding in the reference population.

**Table 3.1** Coefficients of additive and dominance components of variance in the covariances of non-inbred relatives

<table>
<thead>
<tr>
<th>Relatives</th>
<th>( \theta_1 )</th>
<th>( \theta_2 )</th>
<th>( \varphi )</th>
<th>( \varphi' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent–offspring(^a)</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Half-sibs</td>
<td>( (1 + F)/4 )</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
</tr>
<tr>
<td>Full-sibs</td>
<td>( (1 + F)/2 )</td>
<td>((1 + F)/2)^2</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>Uncle–nephew</td>
<td>( (1 + F)/4 )</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
</tr>
<tr>
<td>Half-uncle–nephew</td>
<td>( (1 + F)/8 )</td>
<td>0</td>
<td>( \frac{1}{4} )</td>
<td>0</td>
</tr>
<tr>
<td>First cousins</td>
<td>( (1 + F)/8 )</td>
<td>0</td>
<td>( \frac{1}{4} )</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) For non-inbred parent because it is one of the relatives
Source: Adapted from Cockerham (1963); Kempthorne (1954)

### 3.3.2 Inbred Relatives

The most common procedure involving inbreeding in maize is continuous selfing starting from a non-inbred reference population \( S_0 \). The \( S_0 \) plants are selfed, giving rise to \( S_1 \) families. Selfing is continued and pedigree maintained so that in each generation of inbreeding the relatives can be traced back to a common parent in any previous generation. Actually, this is only one specific situation. In general, we can start in any generation \( t \) of inbreeding and continue selfing to the \( t' \) and \( t'' \) generations. Usually \( t' = t + 1 \) and \( t'' = t' + 1 \), but not necessarily (Cockerham, 1963). As explained earlier two situations exist relative to the gene frequency and the genetic model.

#### 3.3.2.1 Gene Frequency is One-Half (Specific Situation)

When gene frequency is assumed to be one-half, the covariance between relatives in the same generation of inbreeding may be expressed as a linear function of the
components of genetic variance from the non-inbred population. The coefficients are

\[ \theta_1 = (1 + F_t) \quad \text{and} \quad \theta_2 = \left[ (1 + F_t) / (1 - F_t) \right] \left( 1 - F_g \right)^2 \]

where \( F_t \) is the inbreeding coefficient of the last common parent in the selfing chain and \( F_g \) is the inbreeding coefficient of progenies or relatives. The coefficient for additive effects \( \theta_1 \) depends only on the inbreeding of the common parent and is not affected by the inbreeding of progenies or relatives. On the other hand, \( \theta_2 \) depends on the inbreeding of both parents \( F_t \) and progenies \( F_g \). For example, let us consider continuous selfing starting from a non-inbred population \( S_0 \).

As an illustration we will find the genetic variance expressed among \( S_2 \) families. It has already been noted that variance among groups of relatives may be expressed as covariance between relatives within groups. The total genetic variance among \( S_2 \) families, therefore, includes the variance among \( S_2 \) families tracing back to the same \( S_0 \) plant and the variance among groups of \( S_2 \) families tracing back to different \( S_0 \) plants, i.e., \( \hat{\sigma}^2_{S_2} \) (total) = \( \hat{\sigma}^2_{S_1/S_2} + \hat{\sigma}^2_{S_0/S_2} \), where \( S_0 \) and \( S_1 \) are taken in terms of \( S_2 \) data. This is similar to the nested mating design, which is described in Chapter 4.

We have \( \hat{\sigma}^2_{S_0} = \hat{\text{Cov}}_{tgg} = \hat{\text{Cov}} \ 022 \), which is the covariance between individuals in the \( S_2 \) generation whose last common parent is one \( S_0 \) plant. In the selfing series given earlier this corresponds to the covariance between individuals that have the same \( i \) but different \( j \) in the general notation. The covariance is

\[ \hat{\text{Cov}} \ 022 = (1 + 0)\hat{\sigma}^2_A + [(1 + 0)/(1 - 0)](1 - 3/4)\hat{\sigma}^2_D = \hat{\sigma}^2_A + (3/4)\hat{\sigma}^2_D \]

The other term \( \hat{\sigma}^2_{S_1/S_2} = \hat{\text{Cov}}_{t'gg} - \hat{\text{Cov}}_{tgg} = \hat{\text{Cov}} \ 122 - \hat{\text{Cov}} \ 022 \), where \( \hat{\text{Cov}} \ 122 \) is the covariance between individuals in the \( S_2 \) generation whose last common parent is in the \( S_1 \) generation. It can be seen that this is the covariance between individuals that contain the same \( ij \) but different \( k \) in the general notation. Thus we have

\[ \hat{\text{Cov}} \ 122 = (1 + \frac{1}{2})\hat{\sigma}^2_A + [(1 + \frac{1}{2})/(1 - \frac{1}{2})](1 - 3/4)\hat{\sigma}^2_D = (\hat{\sigma}^2_A + (3/8)\hat{\sigma}^2_D \]

and \( \hat{\text{Cov}} \ 122 - \hat{\text{Cov}} \ 022 = (\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{16})\hat{\sigma}^2_D \)
Hence the total genetic variance among $S_2$ families is
\[
\hat{\sigma}^2_{S_2} = \hat{\sigma}^2_A + (\gamma_1)\hat{\sigma}^2_D + (\gamma_2)\hat{\sigma}^2_A + (\gamma_3)\hat{\sigma}^2_D = \hat{\sigma}^2_A + (\gamma_1)\hat{\sigma}^2_D.
\]

The total genetic variance in the $S_2$ generation when gene frequency is one-half is
\[
(1 + F_g)\hat{\sigma}^2_A + (1 - F_g^2)\hat{\sigma}^2_D = (\gamma_4)\hat{\sigma}^2_A + (\gamma_1)\hat{\sigma}^2_D
\]
so $\hat{\sigma}^2$ (total) $- \hat{\sigma}^2$ (among lines) $= (\gamma_4)\hat{\sigma}^2_A + (\gamma_1)\hat{\sigma}^2_D$ is expected to be the genetic variance within $S_2$ families over the population of families.

When relatives are in different generations of inbreeding ($g$ and $g'$), the coefficients are $\theta_1 = 1 + F_t$ and $\theta_2 = [(1 + F_t)/(1 - F_t)](1 - F_g)(1 - F_g')$, which includes the situation when inbreds are in the same generation ($g = g'$). However, when relatives are in different generations the covariance can only be translated into components of genetic variance when there are no additive by dominance epistasis and no dominance types of epistasis (Cockerham, 1963). For example, the covariance between parent and offspring when parents are in the $S_0$ generation is $\hat{\text{Cov}}_{011} = \hat{\sigma}^2_A + (\gamma_2)\hat{\sigma}^2_D + \hat{\sigma}^2_{AA} + \hat{\sigma}^2_{AAA} + \cdots$ if we assume no epistasis involving dominance effects.

### 3.3.2.2 Gene Frequency is Arbitrary (General Situation)

In this case covariance between relatives can only be linearly expressed as a function of the components of variance under the restriction of the genetic model. The only adequate model is one including additive and additive types of epistatic effects. There is, therefore, only one coefficient, $\theta_1 = 1 + F_t$, and the components of variance are defined for the non-inbred generation. Even under dominance effects we can have a good approximation to the additive model with low inbreeding of parents ($F_t \approx 0$) or high inbreeding of progenies or relatives ($1 - F_g \approx 0$) as shown by Cockerham (1963).

Cockerham presented the general limitations of expressing covariances of relatives in linear functions of components of genetic variance, which are summarized in Table 3.2.

When $F = 0$ no limitation is imposed either on the genetic model or on gene frequencies.

With partially inbred relatives ($F < 1$) in the same generation, the expression of covariances as linear functions of genetic variances depends on assumptions about either gene frequency or genetic model. For example, the covariance among $S_1$ families falls into this category because $F = \frac{1}{2}$. Thus for arbitrary gene frequency (unspecified), $\hat{\text{Cov}}_{S_1} = \hat{\sigma}^2_A + \text{all additive types of epistasis}$. When only additive effects are assumed (additive model), it is equal to the variance among $S_1$ families as already shown. If gene frequency is one-half, the covariance can also be linearly related to dominance variance and $\hat{\text{Cov}}_{S_1} = \hat{\sigma}^2 + (\gamma_1)\hat{\sigma}^2_D + \text{all types of epistasis} [\hat{\sigma}^2_{AA} + (\gamma_1)\hat{\sigma}^2_{DD} + \cdots]$. 
Table 3.2  Situations for which the covariance of relatives can be expressed as a linear function of components of genetic variance (Cockerham, 1963)

<table>
<thead>
<tr>
<th>Relatives</th>
<th>Inbreeding coefficient</th>
<th>Gene frequencies</th>
<th>Genetic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-inbreda</td>
<td>( F = 0 )</td>
<td>Unspecified</td>
<td>Unlimited</td>
</tr>
<tr>
<td>In the same generation of inbreeding</td>
<td>( 0 &lt; F &lt; 1 )</td>
<td>Unspecified</td>
<td>Additive and all additive types of epistasis</td>
</tr>
<tr>
<td></td>
<td>( F = 1 )</td>
<td>Unspecified</td>
<td>Unlimited but including only additive and all additive types of epistasis</td>
</tr>
<tr>
<td>In different generations of inbreeding</td>
<td>( 0 \leq F \leq 1 )</td>
<td>One-half</td>
<td>Unlimited</td>
</tr>
<tr>
<td></td>
<td>One-half</td>
<td>Unspecified</td>
<td>Additive and all additive types of epistasis</td>
</tr>
</tbody>
</table>

\( a \) From either non-inbred or inbred reference populations

When relatives are completely homozygous (\( F = 1 \)), there is no limitation on either gene frequencies or genetic model. There is no dominance variance, however, because there are no heterozygotes; consequently only additive and additive types of epistasis effects are considered. Completely inbred relatives can only be obtained by selfing, so the last common parent must have \( F_t = 1 \) and then \( \hat{\text{Cov}}_{tgg} = 2\hat{\sigma}_A^2 \). For the selfing series the approach to homozygosity is theoretically asymptotic and we can say that the limit of \( F_g \) is unity when \( n \), the number of selfing generation, increases toward infinity.

When relatives are in different generations of inbreeding, it is required that the components of genetic variance be translated from one generation to another. This is only valid, assuming either a completely additive model or if dominance effects are not important, because of high inbreeding level of the relatives. If gene frequencies are assumed to be one-half, restriction on the genetic model must be only in the absence of epistasis involving dominance effects.

References


Average allele frequency at segregating loci of F$_2$ populations derived from pure line crosses is expected to be $p = q = 0.5$. From now on, however, we will use as source material genetic broad-based populations with arbitrary allele frequencies. This means that special case of $p = q = 0.5$ often does not apply. Therefore, $p$ is not equal to $q$.

The estimation of variance components in genetically broad-based populations provides the breeder information on their genetic structure. As a consequence, the breeder can select breeding methodologies that are adequate to each population based on the genetic information gathered. Our objective is to calculate several parameters from these populations such as $\hat{\sigma}^2_G = \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_I$ (if possible), $h^2$, $\Delta G$ for each trait of interest. Mating designs generate progenies that are evaluated for the estimation of components of variance. These progenies involve relationships among relatives having known genetic components of variance. For instance, the genetic expectations for HS, PO, and FS covariances are $(\frac{1}{4})\hat{\sigma}^2_A$, $(\frac{1}{2})\hat{\sigma}^2_A$, and $(\frac{1}{2})\hat{\sigma}^2_A + (\frac{1}{4})\hat{\sigma}^2_D$, respectively (see Chapter 3). In addition, progenies should be evaluated over environments with appropriate experimental designs in order to estimate not only components of genetic variance but also components of environmental variance. Experimental designs are analyzed (ANOVA) to estimate experimental error (multi-location replicated trials) and obtain unbiased results (e.g., randomization). Thus expectations (expected mean squares or E(MS)) are expressed in terms of components of variance. These components of variance are then translated to the covariance of relatives depending on the mating design utilized. Finally, we can estimate genetic parameters and make inferences about populations.

It is important to start with good germplasm. Population inferences help in making that decision within breeding programs. As seen before, a population with the best mean is a good choice since genetic gain through years is relatively similar across populations. It is important to analyze representative samples from reference populations both at the phenotypic and molecular level in order to make accurate inferences (e.g., the NAM population recently developed might not be a good example, in a similar way B73 genome does not represent the unique alleles in the maize genome).

Information of the total phenotypic variation that is conditioned by the joint action of genetic and environmental forces is very important for the breeder in
making decisions for the allocation of resources and expected response to selection. Plant breeders observe and measure phenotypes that are the expression of genotypes in a particular set of environments; i.e., \( p_i = u + g_i + e_i + (ge)_i \), where the phenotype \( p_i \) is the sum of an overall mean \( u \), a genotypic effect \( g_i \), an environmental effect \( e_i \), and an interaction effect of genotypes with environments \((ge)_i\). This linear summation of effects shows that units of measure \( p_i \) include not only genotypic effects but also environmental effects on genotypes, which may or may not be repeatable from one environment to another. The environmental ‘noise’ is a very important parameter in maize breeding because it is often beyond the control of the breeder.

In some instances some of the macroenvironmental factors can be controlled (e.g., supplemental moisture by irrigation; application of fertilizers, herbicides, and insecticides; and tillage methods). Genotypes (e.g., inbreds and hybrids) can be screened for drought tolerance under controlled conditions (Carena et al., 2009a) or new cultivars can be developed under controlled low N2 levels if N2 × hybrid interactions are present. In other cases, maize cultivars for reduced tillage conditions can be developed under conventional tillage conditions due to the lack of tillage × hybrid interactions (Carena et al., 2009b). Moreover, breeding programs might not need to develop cultivars under organic conditions if interactions are absent. For some factors we have previous information (e.g., soil types, previous cropping patterns, and long-time weather records); some are uncontrollable because of the uncertainties of the weather (e.g., cloud cover, occurrence of heavy rains and high temperatures, and in temperate areas the duration of frost-free days) and ravages of insects and diseases, which also are influenced by environmental conditions. Hence we measure a phenotype that includes not only the genotypic effect but the influences of an unestimable number of environmental factors on the genotype that are regular and irregular and predictable and unpredictable. Additionally, micro-environmental factors are often more obscure than macroenvironmental factors; these variations occur because of minor variations of the treatments applied to the experimental area, unevenness of moisture flow and penetration, and husbandry practices that may cause minor plant damage. Because the phenotype is a joint expression of the genotypic and environmental effects, our main interest is to determine what
proportion of the phenotypic expression is due to genotypic and environmental effects.

As Cockerham (1956a) has emphasized, the genotypic effect for a particular genotype is the difference between the mean of all the phenotypes with that genotype and the mean of all the phenotypes in the population. Hence the genotypic effect is defined only in contrast with other genotypes in the same environments. If the genotypes vary in expression from one environment to another, the relative contrasts among genotypes can either remain the same or change in relative magnitude and sign. Because the contrasts among genotypes may change, and evidence shows that they often do, it becomes necessary to evaluate genotypes in different environments to determine their general performance. The change in order, ranking, and relative values among genotypes over several environments is the genotype–environment interaction, which is due mostly to macroenvironmental effects.

Properly designed experiments repeated over environments will permit estimation of the variability due to environmental effects and determine its relative importance for the particular genotypes evaluated for the particular sets of environments tested. The total variability for the linear model becomes \( \hat{\sigma}^2_p = \hat{\sigma}^2_g + \hat{\sigma}^2_e + \hat{\sigma}^2_{ge} \), assuming no correlation of genotypes and environments. Proper experimental designs and randomization procedures will tend to minimize the correlations of genotypes and environments, and the additive model allows the estimation of components of variance.

We are basically concerned with variance components estimation in this chapter. Estimation is based on the analysis of variance for balanced data and on expected mean squares proper for the model. In estimating components of genetic variance, a random model is assumed and a random sample of the material for analysis is always required. The estimation procedure involves two steps (Searle, 1971):

1. In the analysis of variance (ANOVA) appropriate to the model, observed mean squares are equal to their expected values; the expected values are linear functions of the unknown variance components, so the resulting equations will be a set of simultaneous linear equations in the variance components.
2. Solve the equations established above; the solutions are the estimators of the variance components.

The expected values of mean squares in the ANOVA do not need assumptions of normality because the variance component estimators obtained by the analysis of variance method of estimation do not, of themselves, depend on normality assumptions. However, the resulting estimators have limited properties (Searle, 1971):

*Unbiasedness.* Estimators of variance components derived by the ANOVA method from balanced data are always unbiased, whether the model is random or mixed.

*Minimum variance.* Variance component estimators obtained by the ANOVA method are minimum variance quadratic unbiased. This means that among
all estimators of $\hat{\sigma}^2$ that are both quadratic functions of the observations and unbiased, those derived by the ANOVA have the smallest variance.

Negative estimates. Variance components are, by definition, positive. Despite this, estimates obtained by the ANOVA method can be negative. Searle (1971) gives some suggestions to overcome this situation. In maize, it seems, negative estimates may be due to an inadequate model (genetic designs to estimate epistatic variance), inadequate sampling (small numbers), and inadequate experimental techniques (competition among progenies).

For estimation of components of variance, we will consider mating designs that develop progenies for evaluation. All mating designs include progenies that involve relationships among relatives having known genetic components of variance (see Chapter 3). For purposes of estimating components of genetic and environmental variance, the progenies developed from the mating design need to be evaluated over environments in an appropriate experimental design. From the analysis of variance of experimental designs, expectations are expressed in terms of the appropriate components of variance; from the components of variance, translations are made to the appropriate relationships (covariances) of relatives based on the mating design used; finally, translations are made from the relationships of relatives to the theoretically determined functions of genetic components of variance for the covariances of relatives.

Unless specified for particular situations, assumptions are necessary for the adequate interpretation of the genetic composition of covariance of relatives across mating designs. The population sampled should have the following:

1) Normal Mendelian diploid inheritance
2) No maternal effects
3) Linkage equilibrium
4) Non-inbred relatives
5) Random selection of parents and relatives
6) No correlation of environmental effects with relatives
7) No epistasis
8) Arbitrary allelic frequencies

Maize generally exhibits regular diploid inheritance, and maternal effects are often not important for most traits. Proper use of experimental design and randomization will ensure no correlation of environmental effects with relatives. Non-inbred relatives are evaluated in most instances, but either inbred or non-inbred parents can be used to produce the relatives as long as the parents are unrelated; in each instance, the parents must have originated from a common population. Because we wish to make inferences from our estimates of genetic components of variance, it is necessary that the relatives are random members of some reference population. The one remaining assumption of linkage equilibrium will probably cause more problems in maize than the others. For certain types of populations (e.g., $F_2$
and recently formed synthetics from inbred lines), linkages could be important in biasing the estimates of components of variance. The possible effects of linkages will be indicated for the different types of populations studied. Random mating of populations before analyses, though, is encouraged.

Some mating designs are used more extensively than others, but each has its own advantages and disadvantages depending on the reference population under consideration and the information desired.

Some generalities to consider are as follows:

---

**Model I vs. model II.** Analyses of variance can be made for fixed (model I) or random (model II) models. In the fixed model, parents of the progenies are the genotypes considered and information gathered is limited to this fixed set of genotypes. On the other hand, parents are a sample of genotypes representing a reference population for the random model.

**Experimental designs.** Incomplete block designs (e.g., lattice) can reduce experimental error. Alpha lattices (Patterson and Williams, 1976) allow more flexibility than lattice designs by having a number of entries that falls between the numbers allowed by lattice designs. Set designs are another type of incomplete block designs that have been used when estimating the variation among entries is more important than obtaining the best estimates of the entry means. The variance components are estimated independently within each set and the estimates are pooled over sets. Each set will give reliable estimates of variance components. Therefore, they can be pooled into one better estimate at the cost of a few degrees of freedom from the error estimate. If the same experiment is grown in a randomized complete block design (RCBD), the number of plots is the same but the only estimate obtained is subjected to the within-block variation. Two types of set designs are used: Reps in Sets and Sets in Reps. Estimates for the genotypic variation are expected to be the same across set designs. However, if in addition to estimating variance components we want to select superior genotype differences among entries from different sets are not precisely compared. Therefore, the set designs might be ideal for estimation of variance components but other designs (e.g., lattice) are better for genotype selection. For instance, we could do selection within sets since we expect to have the same number of superior entries if genotypes are chosen at random across sets. Another possibility could be to compare genotypes across sets by adjusting entry means based on common checks across sets. These designs assume that we have enough field space, labor, and seed supply. However, we often have limited seed supply of numerous early-generation lines for testing. Therefore, there are some designs that allow adequate entry comparisons with fewer replications. The replicated check design includes a single randomized plot of each genotype and check plots of known genotypes at regular intervals. The error variance is based on the variation among checks. The augmented design (Federer, 1961) is similar to the previous design but intends to reduce environmental effects in order to produce unbiased results. The experimental area is divided into blocks and plots are within blocks. Two or more checks are randomly assigned to plots within blocks with the same checks used across blocks. A different set of genotypes is added to each block but a common set
of checks is replicated over blocks. Adjustment is made based on differences among blocks. The replicated checks permit the estimation of $\hat{\sigma}^2$. Unreplicated experiments across locations are very common for early-generation hybrid trials.

**Cross-classification vs. nesting.** A treatment design can be factorial or nested. However, a treatment design does not give the information that experimental designs can give (e.g., replications). Therefore, a treatment design only shows how treatments are defined and it is additional information for the experimental design (e.g., a factorial treatment design in a lattice experimental design). A complete factorial treatment design has all factors cross-classified (e.g., all levels of factor A correspond specifically to equivalent levels of factor B). In the example below, two mating designs (design I and design III) are compared for this purpose. For design I we would say females are nested within males, because each male has a different set of females. Therefore, the levels of one factor do not correspond to levels of the other one. On the other hand, design III would be cross-classified.

<table>
<thead>
<tr>
<th>Design III</th>
<th>Design I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parents</strong></td>
<td><strong>Females</strong></td>
</tr>
<tr>
<td>$P_1$</td>
<td>$f_1$</td>
</tr>
<tr>
<td>$P_2$</td>
<td>$f_4$</td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
</tr>
<tr>
<td>$m_1$</td>
<td>x</td>
</tr>
<tr>
<td>$m_2$</td>
<td>x</td>
</tr>
<tr>
<td>$m_3$</td>
<td>x</td>
</tr>
<tr>
<td>$m_4$</td>
<td>x</td>
</tr>
<tr>
<td>$m_5$</td>
<td>x</td>
</tr>
</tbody>
</table>

Further information on mating designs in plants is given by Cockerham (1963). How good is the theory? Close enough to make progress in selection.

### 4.1 Bi-parental Progenies

Mather (1949) gave this terminology to one of the simplest mating designs for estimation of genetic variance in a reference population. *This type of mating design generates preliminary information of the amount of variability present with minimum effort and cost (e.g., cross-pollinated species).* It provides information needed to determine if significant variation is present in a population for a long-term selection program but no information is available for the type of genetic variation.

This mating design involves crossing pairs of individual plants taken at random from a population. For maize, individual pairs of plants can be crossed reciprocally to produce two ears, which can be bulked for evaluation across environments. To permit adequate interpretations relative to the reference population, it is desirable that one make as many crosses (e.g., a large sample of full-sib families) as facilities permit for growing and making measurements.
If $n$ plants are chosen there will be $n/2$ crosses.

\[
(P1 \times P2) \quad (P3 \times P4) \quad (P5 \times P6) \quad \ldots \quad (P399 \times P400)
\]

\[
FS_1 \quad FS_2 \quad FS_3 \quad \ldots \quad FS_{200}
\]

\[\hat{\sigma}_c^2 = \text{variance among crosses}\]

The information estimable from the among- and within-cross sources of variation are shown in Table 4.1 for the degrees of freedom $df$, mean squares $MS$, and expected mean squares $E(MS)$. An $F$-test of differences among crosses can be made to determine if they are greater than within-cross variations. It is also recognized that an intraclass correlation $r_I$ can be calculated from the analysis of variance in Table 4.1:

\[r_I = \frac{\hat{\sigma}_c^2}{\hat{\sigma}_c^2 + \hat{\sigma}_w^2}\]

Table 4.1 Analysis of variance among- and within-bi-parental crosses

<table>
<thead>
<tr>
<th>Source</th>
<th>$df^a$</th>
<th>MS</th>
<th>$E(MS)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among crosses</td>
<td>$(n/2) - 1$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}_w^2 + k\hat{\sigma}_c^2$</td>
</tr>
<tr>
<td>Within crosses</td>
<td>$(n/2)(k - 1)$</td>
<td>$M_1$</td>
<td>$\hat{\sigma}_w^2$</td>
</tr>
<tr>
<td>Total</td>
<td>$(nk/2) - 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$a^n$ and $k$ refer to the number of parents sampled and plants within each cross, respectively

Mather (1949) presented genetic expectations for among- and within-bi-parental progenies when $F = 0$. Since any among-group variance component is equal to the covariance of the individuals within the groups the variance among crosses ($\hat{\sigma}_c^2$) in Table 4.1 is equal to the $\hat{\text{Cov}}$ FS. In other words, the more alike (covariance depends on resemblance) are individuals within progenies the more different are among progenies.

\[
\hat{\sigma}_c^2 = \hat{\text{Cov}} FS = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{1}{4})\hat{\sigma}_D^2
\]

\[
\hat{\sigma}_c^2 = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{3}{4})\hat{\sigma}_D^2 + \hat{\sigma}_w^2
\]

And the variance within crosses is

\[
\hat{\sigma}_w^2 = [\hat{\sigma}_G^2 - \hat{\text{Cov}} FS] + \hat{\sigma}_w^2
\]

\[
= (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{3}{4})\hat{\sigma}_D^2 + \hat{\sigma}_w^2
\]

$\hat{\sigma}_w^2$ is the environmental source of variation for the variance within crosses. The genetic component $[\hat{\sigma}_G^2 - \hat{\text{Cov}} FS]$ is often described as $\hat{\sigma}_wg^2$.

This type of mating design provides information needed to determine if significant genetic variation is present in a population but no information is available for the type of genetic variation. Growing bi-parental progenies in replicated trials will give the analysis of variance shown in Table 4.2.
Similar to Table 4.1, an F-test is used to determine if the variation among crosses is significantly different from zero, and an intraclass correlation can be computed. If one desires an alternative to the F-test for testing the variation among crosses, a chi-square test can be calculated by dividing the among-cross sum of squares by the error mean square.

Mather (1949) has presented the genetic expectations for the among- and within-bi-parental progenies. The cross component of variance, \( \hat{\sigma}_c^2 \), is the variance among bi-parental progeny means, which is \( \hat{\sigma}_c^2 = \frac{1}{2} \hat{\sigma}_A^2 + \frac{1}{4} \hat{\sigma}_D^2 \) [i.e., \( \frac{1}{4}D + \frac{1}{16}H \) in Mather’s notation], or the \( \hat{\text{Cov}} \) FS. The within-cross component, \( \hat{\sigma}_w^2 \), is the mean variance of bi-parental progenies and includes genetic (\( \hat{\sigma}_{wg}^2 \)) and environmental (\( \hat{\sigma}_{we}^2 \)) sources of variation. The genetic variation within crosses is the total genetic variance, \( \hat{\sigma}_G^2 \), minus the \( \hat{\text{Cov}} \) FS: \( \hat{\sigma}_G^2 - \hat{\text{Cov}} \) FS = \( \frac{1}{2} \hat{\sigma}_A^2 + \frac{1}{4} \hat{\sigma}_D^2 \) [i.e., \( \frac{1}{4}D + \frac{1}{16}H \) in Mather’s notation]. Hence, we have four unknowns (\( \hat{\sigma}_A^2, \hat{\sigma}_D^2, \hat{\sigma}_p^2, \) and \( \hat{\sigma}_{we}^2 \)) and only three mean squares. An estimate of \( \hat{\sigma}_p^2 \) can be obtained from the experimental error mean square (\( M_2 \)), but an estimate of \( \hat{\sigma}_{we}^2 \) is not available from the analysis of variance. If we make the assumption that the dominance effects are zero, we can, however, obtain an estimate of heritability (\( \hat{h}^2 \)). The expressions reduce to \( \hat{\sigma}_c^2 = \frac{1}{2} \hat{\sigma}_A^2 \) and \( \hat{\sigma}_w^2 = \frac{1}{2} \hat{\sigma}_A^2 + \hat{\sigma}_{we}^2 \). Both components of variance can be estimated from the expected mean squares listed in Table 4.2.

Therefore, estimates of \( \hat{\sigma}_A^2, \hat{\sigma}_p^2, \) and \( \hat{\sigma}_{we}^2 \) can be obtained. An estimate of \( \hat{\sigma}_G^2 \) is available from the ANOVA (experimental error mean square) as

\[
\hat{\sigma}_p^2 = \frac{(M_2 - \text{MS}_1)}{k}
\]

If we assume that \( \hat{\sigma}_D^2 = 0 \) we can estimate \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_{we}^2 \) as well as \( \hat{h}^2 \). Thus,

\[
\hat{\sigma}_c^2 = \frac{1}{2} \hat{\sigma}_A^2. \quad \text{Hence,} \quad \hat{\sigma}_A^2 = 2\hat{\sigma}_c^2 \quad \text{and} \quad \hat{\sigma}_{we}^2 = (\hat{\sigma}_w^2 - \hat{\sigma}_c^2).
\]

Again, we can estimate components of variance from expected mean squares of the ANOVA:
\[ \hat{\sigma}^2_A = 2(M_3 - M_2)/rk \]

From the estimates of \( \hat{\sigma}^2_A, \hat{\sigma}^2_p, \) and \( \hat{\sigma}^2_{we}, \) one can determine what proportion of the total variation, for the assumptions used, in the population is under additive genetic control:

\[ \hat{h}^2 = \frac{\hat{\sigma}^2_A}{\hat{\sigma}^2_A + \hat{\sigma}^2_p + \hat{\sigma}^2_{we}}. \]

The approximation becomes more accurate as \( \hat{\sigma}^2_D \) approaches zero.

We can consider an approximation of the population variation under genetic control:

\[ \hat{h}^2 = \frac{\hat{\sigma}^2_c}{\hat{\sigma}^2_w/\hat{\sigma}^2_p + \hat{\sigma}^2_p/r + \hat{\sigma}^2_c} \]  

Based on FS family means

\[ \hat{h}^2 = \frac{(\lambda^2)\hat{\sigma}^2_A}{\hat{\sigma}^2_w/\hat{\sigma}^2_p + \hat{\sigma}^2_p/r + \hat{\sigma}^2_c} \]  

Based on FS family means

If individual plant data are not collected, the estimate of heritability among cross or full-sib family means would be similar: \( \hat{\sigma}^2_c/(\hat{\sigma}^2_p/r + \hat{\sigma}^2_c). \) Confidence limits for the estimates of heritability can be calculated as defined by Knapp et al. (1985).

Standard errors (SE) are associated to variance estimates and should be considered. A variance of the estimate of \( \hat{h}^2 \) can be calculated because the components were calculated from linear functions of independent mean squares.

The variance of a ratio \( \theta \) is approximately:

\[ V(\theta) = V(X)Y^2 - 2[X\widehat{\text{Cov}}(X,Y)]/Y^3(X^2/Y^4)V(Y) \]

where \( (\hat{\sigma}^2_A) = X \) and \( \hat{\sigma}^2_A + \hat{\sigma}^2_p + \hat{\sigma}^2_{we} = Y \) (Kempthorne, 1957). Snedecor (1956, p. 262) has shown that if the variance component was computed from a linear function of independent mean squares, the approximate variance \( V \) of \( \hat{\sigma}^2_i \) is determined as

\[ V(\hat{\sigma}^2_i) = (2/f^2) \Sigma_i (\lambda_i^2 M_i^2)/(df_i + 2) \]

where \( \lambda_i = \pm 1 \) and \( M_i \) are the mean squares used to determine the component of variance, \( df_i \) is the degrees of freedom of the respective mean squares, and \( f \) is the coefficient of the component of variance. Since \( \hat{\sigma}^2_c \) was calculated as \( (M_3 - M_2)/(rk) \), the variance of \( \hat{\sigma}^2_c \) is

\[ V(\hat{\sigma}^2_c) = \frac{2}{(rk)^2} \left[ \frac{M_3^2}{(n/2) + 1} + \frac{M_2^2}{(r - 1)(n/2) + 1} \right] \]

A satisfactory but conservative approximation to the SE of heritability that is easier to calculate was suggested by Dickerson (1969). If we estimate heritability as follows:
\[
\hat{h}^2 = \frac{2\hat{\sigma}_c^2}{(2\hat{\sigma}_c^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{we}^2)}
\]

then,

\[
\text{SE}(\hat{h}^2) = 2\text{SE}(\hat{\sigma}_c^2)(2\hat{\sigma}_c^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{we}^2)
\]

where \(\text{SE}(\hat{\sigma}_c^2)\) is the square root of the variance of \(\hat{\sigma}_c^2\) and the denominator is the phenotypic variance (Knapp et al., 1985; Knapp 1986).

The use of bi-parental progenies provides the breeder with limited information on the relative importance of additive genetic variance. It is a simple mating design to use, but the information may not be sufficient to formulate a long-term breeding program. The information provided by bi-parental progenies gives an indication if indeed sufficient genetic variability is present in the population to warrant a selection program.

### 4.2 Pure Line Progenies (Analysis in Self-Pollinated Crops)

The assumption of no dominance can be reliable in self-pollinated crops since it is either absent or in very low values. Therefore, heterosis \((H)\) is not important since following Falconer and Mackay (1996):

\[
H = \sum y^2 d
\]

being \(d = \) dominance and
\(Y^2 = \) difference in allele frequencies between parents

The general procedure is very simple too:

![Diagram](attachment://diagram.png)

\(S_8\) (unselected sample used for analysis)
4.3 Parent–Offspring Regressions

The ANOVA is as follows:

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>$r-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progenies</td>
<td>$g-1$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_g^2$</td>
</tr>
<tr>
<td>Error</td>
<td>$(r-1)(g-1)$</td>
<td>$M_1$</td>
<td>$\hat{\sigma}^2$</td>
</tr>
</tbody>
</table>

Or across environments as follows:

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments</td>
<td>$e-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reps/envs</td>
<td>$e(r-1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progenies</td>
<td>$g-1$</td>
<td>$M_3$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_g^2 + re\hat{\sigma}_e^2$</td>
</tr>
<tr>
<td>Prog × Envs</td>
<td>$(g-1)(e-1)$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_ge$</td>
</tr>
<tr>
<td>Pooled error</td>
<td>$e(r-1)(g-1)$</td>
<td>$M_1$</td>
<td>$\hat{\sigma}^2$</td>
</tr>
</tbody>
</table>

Since with F$_2$ populations we can assume that $p = q = \frac{1}{2}$, then at the S$_8$ level of inbreeding $F \sim 1$ and $\hat{\sigma}_g^2 = \hat{\sigma}_A^2$. Therefore, heritability can be estimated:

\[
\hat{h}^2 = \frac{2\hat{\sigma}_A^2}{\hat{\sigma}^2 / re + r\hat{\sigma}_g^2 / e + \hat{\sigma}_g^2 / e} \quad \text{Narrow sense heritability}
\]

S7 progeny mean basis

Additional information on the use of maize lines can be found in Section 4.10.

4.3 Parent–Offspring Regressions

In maize we can determine estimates of parent–offspring regressions by (1) regression of offspring on one parent (half-sib method), (2) regression of offspring on the mean of two parents (full-sib method), and (3) regression of selfed progeny on parents (selfing method).

Assume a reference population and make individual plant measurements for the traits of interest (e.g., grain yield, days to maturity, fast dry down, grain quality, lodging resistance, drought tolerance). Harvest seed of the measured plants in the population and measure the same traits in the offspring of each parent. Our task is to determine the degree of association between the traits measured in the parents and in their respective offspring. Our reference population is that from which the parents were derived; in maize it could be either a genetically broad-based variety or an F$_2$ population derived from a cross of two inbred lines. Because maize is cross-pollinated, the individual plants measured would need controlled pollinations for parental control.
Robinson et al. (1949) gave an example of the use of parent–offspring in maize. Randomly chosen $S_0$ plants used as males were crossed to another set of randomly chosen $S_0$ plants used as females. Measurements for quantitative traits were made for the plants used as males and females. Progenies from the crosses were evaluated in trials, and the same traits were measured in the replicated progeny trials.

If the reference population is an $F_2$ population, a common procedure is to measure $F_2$ plants and their progeny means. There is some conflict in terminology, but in both instances $S_0$ parental plants and their progeny are evaluated. As defined previously in the book, because the $F_2$ population is the reference population, the individual $F_2$ plants are equivalent to $S_0$ parents in an open-pollinated variety. In both instances we want to minimize selection of parental $S_0$ plants because we want to estimate genetic parameters relative to the reference population – whether it is an $F_2$ of two inbred lines or an open-pollinated, a synthetic, or a composite variety – to provide an unbiased estimate of heritability.

Our analysis is the regression of the $n$ progeny measurements ($Y$, the dependent variable) on the $S_0$ parental plant measurements ($X$, the independent variable). We use the standard regression model of

$$Y_i = a + bX_i + e_i$$

where $Y_i$ is the mean measurement of the offspring, $X_i$ is the measurement of the parental $S_0$ plant, $b$ is the regression of $Y_i$ on $X_i$, and $e_i$ is the error associated with the $Y_i$. We want to find

$$b = \frac{\sum xy}{\sum x^2} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x^2}$$

because $n - 1$ is common to the numerator and denominator. We know that $\sigma_{xy}$ has the following genetic component (covariance of parent–offspring, Cov PO):

$$\hat{\text{Cov PO}} = \sum_i \left(\frac{1+F_i}{2}\right)^i \hat{\sigma}_A^2$$

When $F = 0$,

$$\hat{\text{Cov PO}} = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{1}{4})\hat{\sigma}_{AA}^2 + (\frac{1}{8})\hat{\sigma}_{AAA}^2 + \cdots$$

and the total variation of the parental measurements is $\hat{\sigma}_x^2$. Then, assuming no epistasis,

$$b = (\frac{1}{2})\hat{\sigma}_A^2/\hat{\sigma}_x^2 = \frac{\hat{\sigma}_A^2}{(2\hat{\sigma}_x^2)}$$

To determine the heritability estimate on an individual plant basis of the traits from parent–offspring regression, we can calculate

$$\hat{h}^2 = 2b = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_x^2}$$
A standard error can be calculated rather easily for the heritability estimate of parent–offspring regression. From standard statistical texts the standard error (SE) of the regression coefficient is determined as

\[
\hat{\sigma}_b^2 = \left[ \sum_i y_i^2 - \left( \sum_i x_i y_i \right)^2 / \sum_i x_i^2 \right] / \left[ (n - 2) \sum_i x_i^2 \right] = \frac{\text{average of deviations from regression squared}}{\sum x_i^2}
\]

Hence, SE(\(b\)) = \((\hat{\sigma}^2 / \Sigma x_i^2)^{1/2}\). The standard error of the heritability estimate on an individual basis becomes SE(\(h^2\)) = 2SE(\(b\)).

Parent–offspring regression also can be obtained by evaluating the progeny from the cross of two individuals; for this case we would have the regression of offspring on the mean of the parents. Mean parent–offspring regression also would be amenable for use in maize because of the ease in making controlled crosses between two individuals. Progeny would be produced in the same manner as described for bi-parental progenies. Instead of taking measurements only on offspring, as for bi-parental progenies, we also want to record measurements of traits in each parent used in crosses and regress offspring measurements \(Y\) on those of the mean of the pairs of parents, \(\bar{X}\). If we use

\[
\bar{X} = (X_1 + X_2)/2
\]

then mid-parent–offspring regression, \(\widehat{\text{Cov PO}}\) becomes

\[
b = \hat{\sigma}_{xy} / \hat{\sigma}_x^2
\]

Assuming \(\hat{\sigma}_{X_1}^2 = \hat{\sigma}_{X_2}^2\) and that \(X_1\) and \(X_2\) are uncorrelated,

\[
\hat{\sigma}_X^2 = (l_4)\hat{\sigma}_{X_1}^2 + (l_4)\hat{\sigma}_{X_2}^2 = (l_2)\hat{\sigma}_{X}^2
\]

Thus,

\[
b = (l_2)\hat{\sigma}_A^2 / [(l_2)\hat{\sigma}_X^2] = \hat{\sigma}_A^2 / \hat{\sigma}_X^2 \text{ and } \hat{h}^2 = b
\]

Mid-parent–offspring regression also would have utility for estimation of heritability in dioecious plant species (e.g., hops, hemp, asparagus, date palms, and willow trees). The species need not be dioecious but \(2n\) parents must be measured for mid-parent–offspring regression, whereas only \(n\) parents must be measured for parent-offspring regression. In all instances the reference population from which the parents are derived is random mated and non-inbred. Sufficient sampling and no selection of the parents evaluated should be ensured for the estimates to have meaning relative to the reference population.

In maize it is also convenient to measure the \(S_0\) plants in either a broad genetic base population or an \(F_2\) population and self the \(S_0\) plants to obtain \(S_1\) progenies.
Measurements can be taken on the $S_0$ plants and $S_1$ progenies in replicated trials. Because the $S_0$ plants were selfed, the measurements represent the male and female parents; hence, the $S_0$ plant measurements would be for both parents. Regression of $S_1$ progenies on the $S_0$ plants provides an estimate of

$$\hat{\text{Cov PO}} = \hat{\sigma}_A^2 + \left(\frac{1}{2}\right) \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 \quad (p = q = 0.5, \text{ see Chapter 3}).$$

It is seen that $\hat{\text{Cov PO}}$ for $S_1$ progenies on $S_0$ plants includes dominance and dominance-type epistasis in addition to additive and additive-type epistasis. A broad sense heritability estimate can be determined as

$$\hat{h}^2 = b = \frac{(\hat{\text{Cov PO}})}{\hat{\sigma}_x^2},$$

where $\hat{\sigma}_x^2$ is the phenotypic variability among the $S_0$ plants.

This would be a conservative estimate of $\hat{h}^2$ because the genetic variability among a population of $S_0$ plants would include all the genetic variability ($\hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \cdots$), whereas $\hat{\text{Cov PO}}$ includes only half the dominance variance and fractions of the epistatic sources of variation. If allelic (dominance) interaction effects are assumed minor, the estimates of heritabilities for $\hat{\text{Cov PO}}$ of $S_1$ progenies on $S_0$ plants would be similar to $\hat{\text{Cov PO}}$ of progenies on one parent and $\hat{\text{Cov PO}}$ of progenies on the mean of parents. The estimate of $\hat{\sigma}_A^2$, however, is not exactly the same as in the previous formula unless either $p = q = 0.5$ or dominance is absent.

The progeny of the parent–offspring that are regressed on the mean of the parents also can be analyzed as shown in Table 4.2. We would have two equations that would permit estimation of the variance due to dominance effects:

$$\hat{\text{Cov FS}} = \left(\frac{1}{2}\right) \hat{\sigma}_A^2 + \left(\frac{1}{4}\right) \hat{\sigma}_D^2 \text{ and } \hat{\text{Cov PO}} = \left(\frac{1}{2}\right) \hat{\sigma}_A^2,$$

assuming no epistasis

Hence $4(\hat{\text{Cov FS}} - \hat{\text{Cov PO}})$ is an estimate of $\hat{\sigma}_D^2$. If individual plant data are collected for the progenies that resulted from the cross of two parents, we have three equations that permit estimation of three parameters.

For instance, additive $\times$ additive epistatic variance could be estimated because

$$\hat{\text{Cov FS}} = \left(\frac{1}{2}\right) \hat{\sigma}_A^2 + \left(\frac{1}{4}\right) \hat{\sigma}_D^2 + \left(\frac{1}{4}\right) \hat{\sigma}_{AA}^2$$

$$\hat{\text{Cov PO}} = \left(\frac{1}{2}\right) \hat{\sigma}_A^2 + \hat{\sigma}_{AA}^2$$

and

$$\hat{\sigma}_G^2 - \hat{\text{Cov FS}} = \left(\frac{1}{2}\right) \hat{\sigma}_A^2 + \left(\frac{3}{4}\right) \hat{\sigma}_D^2 + \left(\frac{1}{4}\right) \hat{\sigma}_{AA}^2$$

If non-additive effects are small relative to additive effects, each equation has similar expectations for $\hat{\sigma}_A^2$ and the observed variances would be similar. If non-additive effects are important, the observed component for the within-plot variance would deviate from the other two.
The two commonly used methods of estimating heritability by use of parent–offspring regression \((2b = \hat{h}^2\) and \(b = \hat{h}^2\)) are valid when the parents are non-inbred, as in a random mating population. If the parents are inbred or related, Smith and Kinman (1965) have shown that the previous inbreeding will cause an upward bias in the estimate of heritability. The bias is not generally severe for bisexual populations, but it becomes more important in self-pollinated populations. Smith and Kinman (1965) have shown that the correct estimator for the general case is

\[
\frac{b}{2r_{XY}}
\]

where \(r_{XY}\) is a measure of the relationship between the parent Y and its offspring X (Kempthorne, 1957). Malécot (1948) defined \(r_{XY}\) as the probability that a random gene at a specific locus in X is identical by descent to a random gene at the same locus in Y. Hence the estimation \(b/(2r_{XY}) = \hat{h}^2\) provides for an adjustment for the mating system (bisexual or self-pollination) and for the known level of inbreeding or relationship of the parents.

Smith and Kinman (1965) showed the instances of parent–offspring regressions that may be used to estimate heritability for continuous self-pollinations:

<table>
<thead>
<tr>
<th>Parent–offspring generation</th>
<th>(r_{XY})</th>
<th>(\hat{h}^2 = b/(2r_{XY}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F_1, F_2)</td>
<td>(\frac{1}{2})</td>
<td>(b_{F_2, F_1})</td>
</tr>
<tr>
<td>(F_2, F_3)</td>
<td>(\frac{3}{4})</td>
<td>((\frac{2}{3})b_{F_3, F_2})</td>
</tr>
<tr>
<td>(F_3, F_4)</td>
<td>(\frac{7}{8})</td>
<td>((\frac{3}{7})b_{F_4, F_3})</td>
</tr>
<tr>
<td>(F_4, F_5)</td>
<td>(\frac{15}{16})</td>
<td>((\frac{8}{15})b_{F_5, F_4})</td>
</tr>
<tr>
<td>(F_5, F_6)</td>
<td>(\frac{31}{32})</td>
<td>((\frac{16}{32})b_{F_6, F_5})</td>
</tr>
</tbody>
</table>

It is seen that adjustments for level of inbreeding make the estimates more conservative when we have continuous self-fertilization. The same estimator is valid for bisexual populations except that \(r_{XY}\) also depends on level of inbreeding of the parents \([Y, (F_Y)]\) and degree of relationship between the two parents \((r_{YZ})\). If the two parents are non-inbred and unrelated, \(2r_{XY} = \frac{1}{2}\); but if both parents are identical homozygotes (both \(F_Y\) and \(r_{YZ}\) are equal to 1), \(2r_{XY} = 2\), which is the same as for continuous self-pollination. Ordinarily, the parents are non-inbred and unrelated for most instances when parent–offspring regression is used for estimation of heritability in maize. But self-fertilization can be used to generate different inbred generations by controlled pollination. If some inbred generations are used to determine heritabilities, adjustments are needed to account for the level of inbreeding. Smith and Kinman (1965) have discussed the situations and shown the relations that are necessary to account for the inbreeding and/or relatedness of the parents when using parent–offspring regression for estimating heritabilities.

Analyses of variance (ANOVAs) can be conducted for fixed or random models. Fixed models, also referred as to model I, are based on a fixed set of genotypes. Thus, parents are the sole genotypes considered. On the other hand, random models, also referred as model II, are based on a random set of genotypes, so parents are a sample of genotypes from a reference population.
### 4.4 Design I

Also known as North Carolina design I or B/A design, the design I mating scheme also was introduced by Comstock and Robinson (1948). Except for the diallel it is one of the most used mating designs in maize since it is the easiest design for producing a large number of progenies. This mating design can also be useful in self-pollinated crops with multiple flowers.

Design I is adequate only for estimating genetic components of variance for a reference population. Hence the model II analysis is applied because the parents used in producing progenies for testing are an unselected sample from the reference population.

Assume a random mating population in linkage equilibrium as the reference population. If we randomly choose S₀ plants from this reference population, then we can assign some of them as males and some of them as females. Each male is crossed to a different set of females (independent sample) to produce progenies for evaluation; i.e., m males are each mated to f females to produce mf progenies for evaluation. The genetic structure of the progenies includes full-sibs that have both parents in common and half-sibs that have a male parent in common. Therefore, expected mean squares can be expressed in covariance of relatives.

If each male is mated to four females, we will have the following matings:

```
Parents  Progenies  Genetic structure
m₁ x f₁  m₁f₁  x
m₂ x f₂  m₁f₂  x
m₃ x f₃  m₁f₃  x
m₄ x f₄  m₁f₄  x
```

or,

```
m₁ x f₁ = p₁₁
m₂ x f₅ = p₂₅
...  
mᵢ x fⱼ = pᵢⱼ
```

```
F₂ = p₁₂
F₃ = p₁₃
F₄ = p₁₄
```

```
f₁ = p₂₆
f₇ = p₂₇
f₈ = p₂₈

```

```
Individuals within each $p_{ij}$ progeny are full-sibs. The $p_{ij}$, $p_{ik}$, $p_{il}$, $p_{in}$ progenies are half-sibs because they have a common $i$ male parent. The model for one environment is

\[ Y_{ijk} = u + m_i + f_{ij} + r_k + e_{ijk}, \]

where $u$ is the mean, $m_i$ is the effect of the $i$th male, $f_{ij}$ is the effect of the $j$th female mated to the $i$th male, $r_k$ is the replication effect, and $e_{ijk}$ is the experimental error. Because the mating design is nested, expected mean squares are obtained by a hierarchical type of design. Also, because of the genetic structure of the mating design, expected mean squares can be expressed in the more useful covariances of relatives (Table 4.3).

**Table 4.3** Analysis of variance of the design I mating design for one environment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>E(MS)</th>
<th>Variance component</th>
<th>Covariance of relatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>$r - 1^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>$m - 1$</td>
<td>$M_4$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}<em>{lm}^2 + rf</em>{m}^2$</td>
<td>$\hat{\sigma}^2 + rf_{m}^{\text{Cov HS}}$</td>
</tr>
<tr>
<td>Females (males)</td>
<td>$m(f - 1)$</td>
<td>$M_3$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_{lm}^2$</td>
<td>$\hat{\sigma}^2 + rf_{m}^{\text{Cov HS}}$</td>
</tr>
<tr>
<td>Error</td>
<td>$(r - 1)(mf - 1)$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2$</td>
<td>$\hat{\sigma}^2$</td>
</tr>
<tr>
<td>Total</td>
<td>$rmf - 1$</td>
<td>$M_1^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>$rmf(k - 1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$, $r$, $m$, $f$, and $k$ refer to number of replications, males, females within males, and plant within plots, respectively. SOV = Source of variation. $^b$ $M_1 = \hat{\sigma}_{we}^2 = (\hat{\sigma}_{we}^2 + \hat{\sigma}_{wg}^2) = [\hat{\sigma}_w^2 + (\hat{\sigma}_G^2 - \text{Cov HS})]$.

Hence, $M_2 = \hat{\sigma}^2 = (\hat{\sigma}_{we}^2 + (\hat{\sigma}_G^2 - \text{Cov FS})) / k + \hat{\sigma}_p^2$, where $\hat{\sigma}_p^2$ is the experimental plot error. The same concept used with bi-parental progenies is utilized with design I. The differences among males are equal to the similarities between half-sib families within males ($\hat{\sigma}_{lm}^2 = \text{Cov HS}$) because the better association within groups, the greater differences among groups.

Direct $F$-tests can be made for males and females-within-males mean squares, and males and females-within-males components of variances can be estimated from the appropriate mean squares. Then covariances of relatives can be related to the genetic components of variance. The male component (Cov HS) is the same genetically as the GCA of the diallel and partial diallel and among males and among females of design II (explained later in the chapter). The among-females-within-males component, however, has a different expectation compared to other designs ($\hat{\sigma}_{lm}^2 = \text{Cov FS} - \text{Cov HS}$). For design I each male is mated to a different group of females. Therefore, only one half-sib-ship is included, and the component, females within males, is

\[ \text{Cov FS} - \text{Cov HS} = (\frac{1}{4})\hat{\sigma}_{A}^2 + (\frac{1}{4})\hat{\sigma}_{D}^2 \text{ for } F = 0 \]
As a consequence, we do not have a direct estimate of \( \hat{\sigma}_D^2 \) from the mean squares of design I, which can only be obtained by solving for expectations of components of variance:

\[
\hat{\sigma}_D^2 = 4\hat{\sigma}_{lm}^2 - 4\hat{\sigma}_m^2 = 4[(\hat{\text{Cov}} \ FS - \hat{\text{Cov}} \ HS) - \hat{\text{Cov}} \ HS] = 4[(\lambda_4)\hat{\sigma}_A^2 + (\lambda_4)\hat{\sigma}_D^2 - (\lambda_4)\hat{\sigma}_A^2] = \hat{\sigma}_D^2
\]

Summarizing, the crosses consist of full-sib individuals but half-sib covariance is present within males. Therefore, the variation among females within males follows the same concept and it is defined as the difference between the similarities among FS progenies and the similarities among HS progenies. Since

\[
\hat{\sigma}_m^2 = \hat{\text{Cov}} \ HS \quad \text{and} \quad \hat{\text{Cov}} \ HS = (\lambda_4)\hat{\sigma}_A^2,
\]

and since

\[
\hat{\sigma}_{lm}^2 = \hat{\text{Cov}} \ FS - \hat{\text{Cov}} \ HS \quad \text{and} \quad \hat{\text{Cov}} \ FS - \hat{\text{Cov}} \ HS = (\lambda_4)\hat{\sigma}_A^2 + (\lambda_4)\hat{\sigma}_D^2 - (\lambda_4)\hat{\sigma}_A^2
\]

\[
\hat{\sigma}_{lm}^2 = (\lambda_4)\hat{\sigma}_A^2 + (\lambda_4)\hat{\sigma}_D^2,
\]

For \( F = 0 \) and no epistasis, \( \hat{\sigma}_G^2 - \hat{\text{Cov}} \ FS = (\lambda_4)\hat{\sigma}_A^2 + (\lambda_4)\hat{\sigma}_D^2. \)

The importance of not having a ‘clean’ estimate of \( \hat{\sigma}_D^2 \) is exemplified by the calculation of standard errors of estimates of components of variance. From Table 4.3 the estimated variance of \( \hat{\sigma}_A^2 \) is

\[
V(\hat{\sigma}_A^2) = \frac{16 \times 2}{r^2f^2} \left[ \frac{M_4^2}{m+1} + \frac{M_3^2}{m(f-1)+2} \right]
\]

The estimate of \( \hat{\sigma}_D^2 \) is obtained as \( 4(\hat{\sigma}_{lm}^2 - \hat{\sigma}_m^2) \), where \( \hat{\sigma}_{lm}^2 = (M_3 - M_2)/r \) and \( \hat{\sigma}_m^2 = (M_4 - M_3)/(rf) \). Therefore, the estimate of the variance of \( \hat{\sigma}_D^2 \) is approximately

\[
V(\hat{\sigma}_D^2) = \frac{16 \times 2}{r^2f^2} \left[ \frac{M_4^2}{m+1} + \frac{(f + 1)^2M_3^2}{m(f - 1) + 2} + \frac{f^2M_2^2}{(r-1)(mf - 1) + 2} \right]
\]

Because of the complicated function used to estimate \( \hat{\sigma}_D^2 \), the variance of the estimate usually is quite large as many covariance components are present. The estimate, therefore, is not as accurate as other estimates.

It is essential that the males and the females mated to each male are randomly chosen in the reference population. Any stratification of the individuals selected as males and females could reduce the estimates of \( \hat{\sigma}_m^2 \) and \( \hat{\sigma}_{lm}^2 \). Randomly chosen males mated to a group of females that are similar in date of flowering to the male
could reduce the variability among females within males ($\hat{\sigma}_{fm}^2$). Assortative mating of males with females can be reduced by delayed planting of a portion of the population from which the males are chosen (Lindsey et al., 1962). Early flowering female silks can be held for later pollination by covering the ear shoot to prevent fertilization. Randomly chosen males from the delayed planting can be crossed with the females without regard to vigor, plant size, maturity, etc. It can be argued that this forced mating is not representative of a cross-pollinated population because of forced crossing among physiologically isolated segments of a random mating population, but it does satisfy the assumptions of the genetic model.

Because we are interested in obtaining valid estimates of genetic components of variance of our reference population, we want to include an adequate sample of genotypes from our reference population. If a large number of males are mated with females, the size of the experiment and replication required to include all progenies could be quite large. To reduce replication size and attempt to increase the precision of our experiment, Comstock and Robinson (1948) suggested grouping progenies into sets by males. If 100 male plants are each mated with four females, we have 400 full-sib progenies for testing. To reduce replication size, five sets of 20 males each or 80 full-sib progenies may be considered. Or assume that 10 males each crossed to 4 females are included in one set of 40 full-sib progenies. The final choice of number of progenies to include in a set depends on the experimenter’s knowledge of soil variability and past experience with size of replication for local control of experimental error. After the number of progenies included in a set has been determined, the two alternatives for arranging the progenies in the field are (1) replications within sets or (2) sets (as sub-blocks) within replications. Since this is a large experiment representing a good sample size we can try to solve small experiments within the big one by using the sets. Sets are used to estimate more accurate components of variance with better control of the experimental error. By making analyses of several small experiments and then making a combined analysis of all experiments we increase the precision of estimates:

<table>
<thead>
<tr>
<th>SET 1</th>
<th>SET 2</th>
<th>SET 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rep 1</td>
<td>Rep 2</td>
<td>Rep 1</td>
</tr>
<tr>
<td>SET 4</td>
<td></td>
<td>And so forth...</td>
</tr>
</tbody>
</table>

Therefore, in order to reduce replication size and to increase the precision of the experiment we group progenies into sets by males depending on past plot experience.

Each set is analyzed as shown in Table 4.4 and then pooled across sets for degrees of freedom and sums of squares as shown in Table 4.5. There are only small differences in the distribution of degrees of freedom; therefore, the major difference between the two field arrangements may be in magnitude of error sums of squares. Intuitively, it seems that replications within sets would be a preferable arrangement for local control of experimental error.
### Table 4.4 Analysis of variance of design I experiment pooled over sets in one environment

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>General</th>
<th>Example</th>
<th>Mean squares</th>
<th>Expected mean squares $^{a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sets</td>
<td>$s-1^{b}$</td>
<td>9</td>
<td></td>
<td>$M_4$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}^2_{I\hat{m}} + rf\hat{\sigma}^2_{m}$</td>
</tr>
<tr>
<td>Replications/sets</td>
<td>$(r-1)$</td>
<td>10</td>
<td></td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2$</td>
</tr>
<tr>
<td>Males/sets</td>
<td>$s(m-1)$</td>
<td>90</td>
<td>$M_4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females/males/sets</td>
<td>$sm(f-1)$</td>
<td>300</td>
<td>$M_3$</td>
<td></td>
<td>$\hat{\sigma}^2 + \hat{\sigma}^2_{I\hat{m}}$</td>
</tr>
<tr>
<td>Pooled error</td>
<td>$(smf-1)(r-1)$</td>
<td>390</td>
<td>$M_2$</td>
<td></td>
<td>$\hat{\sigma}^2$</td>
</tr>
<tr>
<td>Total</td>
<td>$srmf-1$</td>
<td>799</td>
<td></td>
<td>$M_1^{c}$</td>
<td></td>
</tr>
<tr>
<td>Within plots</td>
<td>$srmf(k-1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Covariances of relatives are as shown in Table 4.3

$^b$r, s, m, f, and k refer to the number of replications (2), sets (10), males (10), females (4), and number of plants within plots, respectively, in the example

$^c$See Table 4.2

### Table 4.5 Analysis of variance of design I experiments pooled over sets and repeated over environments for sets within replications

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean squares</th>
<th>Expected mean squares $^{a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments (E)</td>
<td>$e - 1^{b}$</td>
<td></td>
<td>$M_6$ $\hat{\sigma}^2 + r\hat{\sigma}^2_{I\hat{m}} + rf\hat{\sigma}^2_{m} + rf\hat{\sigma}^2_{\hat{m}} + rf\hat{\sigma}^2_{m}$</td>
</tr>
<tr>
<td>Replications/E</td>
<td>$(e-1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sets/replications/E</td>
<td>$(e-1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males/sets</td>
<td>$s(m-1)$</td>
<td>$M_6$</td>
<td></td>
</tr>
<tr>
<td>Females/males/sets</td>
<td>$ms(f-1)$</td>
<td>$M_5$</td>
<td></td>
</tr>
<tr>
<td>E × males/sets</td>
<td>$(e-1)s(m-1)$</td>
<td>$M_4$</td>
<td></td>
</tr>
<tr>
<td>E × females/males/sets</td>
<td>$(e-1)ms(f-1)$</td>
<td>$M_3$</td>
<td></td>
</tr>
<tr>
<td>Pooled error</td>
<td>$(e-1)(mf-1)$</td>
<td>$M_2$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$esrnf-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within plots</td>
<td>$esrnf(k-1)$</td>
<td></td>
<td>$M_1^{c}$</td>
</tr>
</tbody>
</table>

$^a$Covariances of relatives are as shown in Table 4.3

$^b$e, r, s, m, f, and k refer to the number of environments, replications with sets, sets, males, females within males, and plants within plots, respectively

$^c$See Table 4.2
The analysis of the design I experiment repeated over environments is shown in Table 4.5. Translation of components of variance to covariances of relatives permits estimation of components of genetic variances (\(\hat{\sigma}^2_A\) and \(\hat{\sigma}^2_D\)) and their interactions with environments.

F-tests and estimation of components of variance can be made directly for all sources of variation except males within sets, for which Satterthwaite’s (1946) approximation can be used.

In summary, design I has been used frequently in maize, and it is good for extensive sampling of S0 plants in a population. In comparison with other mating designs and other crops, design I is the easiest for producing a large number of progenies in maize. The nested structure of the progenies makes design I amenable for grouping them in sets, and the pooling across sets is straightforward. Estimates of additive genetic variance (\(\hat{\sigma}^2_A\)) and total genetic variance (\(\hat{\sigma}^2_G\)), assuming no epistasis, are obtained directly from the mean squares of the analysis of variance. If an estimate of dominance variance (\(\hat{\sigma}^2_D\)) is desired, it can be obtained as the difference between the females-within-males and the males components of variance, but the variance of \(\hat{\sigma}^2_D\), unfortunately, is usually quite large. Similar to diallel and design II analyses, design I provides GCA information for males. If male plants are self-pollinated, early test information is obtained and males with superior GCA can be included in breeding nurseries as S1 progenies. Average dominance of the genes also can be determined from design I analyses. For restrictions of no epistasis, linkage equilibrium, and \(p = q = 0.5\), expectations of the components of variance are

\[
\hat{\sigma}^2_m = \text{Cov HS} = \langle y \rangle \hat{\sigma}^2_A \quad \text{where} \quad \hat{\sigma}^2_A = \langle y \rangle \sum a_i^2
\]

\[
\hat{\sigma}^2_{f/m} = \text{Cov FS} - \text{Cov HS} = \langle y \rangle \hat{\sigma}^2_A + \langle y \rangle \hat{\sigma}^2_D \quad \text{where} \quad \hat{\sigma}^2_D = \langle y \rangle \sum d_i^2
\]

And the average gene dominance is

\[
\bar{d} = \left[ \frac{2(\hat{\sigma}^2_{f/m} - \hat{\sigma}^2_m)}{\hat{\sigma}^2_m} \right]^{\frac{1}{2}} = \left( \frac{2\hat{\sigma}^2_D}{\hat{\sigma}^2_A} \right)^{\frac{1}{2}} = \left[ \frac{\langle y \rangle \sum d_i^2}{\langle y \rangle \sum a_i^2} \right]^{\frac{1}{2}}
\]

Heritability estimates based on the mean of \(r\) plots can be determined from components of variance given in Table 4.4 for one environment as follows:

\[
\hat{h}^2 = \frac{4\hat{\sigma}^2_m}{\hat{\sigma}^2/r + 4\hat{\sigma}^2_{f/m}}
\]

For non-inbred parents and no epistasis, \(\hat{\sigma}^2_m = \langle y \rangle \hat{\sigma}^2_A\) and \(\hat{\sigma}^2_{f/m} = \langle y \rangle \hat{\sigma}^2_A + \langle y \rangle \hat{\sigma}^2_D\); hence, \(4\hat{\sigma}^2_{f/m}\) includes additive and dominance variance. The standard error of this estimate of heritability is approximately \(\text{SE}(\hat{h}^2) = 4\text{SE}\hat{\sigma}^2_m/(\hat{\sigma}^2/r + 4\hat{\sigma}^2_{f/m})\) where \(\text{SE}(\hat{\sigma}^2_m)\) is the square root of
Heritability estimates for individual plant selection can be calculated as \( \hat{h}^2 = \frac{4\hat{\sigma}_m^2}{\hat{\sigma}_w^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{lm}^2 + \hat{\sigma}_m^2} \), where \( \hat{\sigma}_w^2 \) is the estimate of the within-plot variability and \( \hat{\sigma}_p^2 \) is estimate of plot error. Standard error of \( \hat{h}^2 \) can be calculated as

\[
\text{SE}(\hat{h}^2) = \frac{4\text{SE}(\hat{\sigma}_m^2)}{\hat{\sigma}_w^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{lm}^2 + \hat{\sigma}_m^2}
\]

These estimates of heritability are for one environment and would include an unknown bias because of genotype–environment interactions, i.e., \( 4\hat{\sigma}_m^2 = \hat{\sigma}_A^2 + \hat{\sigma}_{AE}^2 \).

From the analysis repeated over environments (Table 4.5) an estimate of heritability unbiased by genotype–environment interactions and based on the mean of \( re \) plots would be

\[
\hat{h}^2 = \frac{4\hat{\sigma}_m^2}{\hat{\sigma}_w^2/(re) + 4\hat{\sigma}_{el/m}^2/e + 4\hat{\sigma}_{lm}^2}
\]

Following Nyquist (1991) and Holland et al. (2003), the estimates of heritability based on half-sib progeny means can be calculated as

\[
\hat{h}^2 = \frac{\hat{\sigma}_m^2}{\hat{\sigma}_w^2/rf + \hat{\sigma}_{lm}^2/f + \hat{\sigma}_m^2} \quad \text{One environment (Table 4.4)}
\]

\[
\hat{h}^2 = \frac{\hat{\sigma}_m^2}{\hat{\sigma}_w^2/ref + \hat{\sigma}_{el/m}^2/ef + \hat{\sigma}_{em}/e + \hat{\sigma}_{lm}^2/f + \hat{\sigma}_m^2} \quad \text{Half-sib progeny means basis}
\]

and as

\[
\hat{h}^2 = \frac{\hat{\sigma}_m^2}{2(\hat{\sigma}_m^2 + \hat{\sigma}_{lm}^2) + 4\hat{\sigma}_{el/m}^2/e + 4\hat{\sigma}_m^2} \quad \text{Across environments (Table 4.5)}
\]

\[
\hat{h}^2 = \frac{\hat{\sigma}_m^2}{2(\hat{\sigma}_m^2 + \hat{\sigma}_{lm}^2) + \hat{\sigma}_w^2/rf + \hat{\sigma}_{lm}^2/f + \hat{\sigma}_m^2} \quad \text{Half-sib progeny means basis}
\]

with confidence limits on the estimates as shown by Knapp et al. (1985). The males are crossed to a series of females to evaluate the combining ability of males and differences among males are because of the covariances of half-sib families. Genetic response to selection would be based on half-sib family means (see Chapter 6).

Non-inbred parents usually have been used in making design I progenies. If, however, parents are either partially inbred or homozygous, the coefficients for the calculation of \( \hat{\sigma}_A^2 \) would have to be changed accordingly, e.g., \( \text{Cov HS} = [(1 + F)/4] \hat{\sigma}_A^2 \). Thus if \( F = 1 \) and \( \hat{\sigma}_m^2 = (\hat{\sigma}_D^2) \hat{\sigma}_A^2 \), all estimates of heritability given are in the narrow sense.

A check on the relative importance of non-additive epistatic effects can be determined by comparing \( 4\hat{\sigma}_m^2 = \hat{\sigma}_A^2 + (\hat{\sigma}_4) \hat{\sigma}_A^2 + \cdots ; 4\hat{\sigma}_{lm}^2 = \hat{\sigma}_D^2 + \hat{\sigma}_4 + (\hat{\sigma}_3) \hat{\sigma}_A^2 + \cdots ; \) and \( 2(\hat{\sigma}_m^2 + \hat{\sigma}_{lm}^2) = \hat{\sigma}_A^2 + (\hat{\sigma}_2) \hat{\sigma}_D^2 + \cdots + (\hat{\sigma}_2) \hat{\sigma}_A^2 + \cdots \). All relations include \( 4\hat{\sigma}_{lm}^2 \).
but they have different proportions of non-additive variance, with $4\hat{\sigma}_m^2$ being the greatest and $4\hat{\sigma}_A^2$ the least bias due to epistasis. Estimates of heritability determined by use of the male component of variance contain less bias due to epistasis than the other two components.

The relative of digenic epistatic effects can be estimated as follows where the amount of epistatic bias depends on the level of inbreeding:

If $F = 0$

$$\hat{\text{Cov}} \text{ HS} = (\frac{1}{4})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AA}^2$$

$$\hat{\text{Cov}} \text{ FS} = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AD}^2 + (\frac{\gamma_4}{8})\hat{\sigma}_{AD}^2 + (\frac{\gamma_4}{16})\hat{\sigma}_{DD}^2$$

If $F = 1$

$$\hat{\text{Cov}} \text{ HS} = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AA}^2$$

$$\hat{\text{Cov}} \text{ FS} = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{DD}^2$$

Therefore, if $F = 0$

$$\hat{\sigma}_m^2 = \hat{\text{Cov}} \text{ HS} = (\frac{1}{4})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AA}^2$$

$$\hat{\sigma}_{mF}^2 = \hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HS} = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AD}^2 + (\frac{\gamma_4}{8})\hat{\sigma}_{AD}^2 + (\frac{\gamma_4}{16})\hat{\sigma}_{DD}^2$$

If $F = 1$

$$\hat{\sigma}_m^2 = \hat{\text{Cov}} \text{ HS} = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{\gamma_4}{4})\hat{\sigma}_{AA}^2$$

$$\hat{\sigma}_{mF}^2 = \hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HS} = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{DD}^2$$

We can, therefore, estimate which component of variance (males or females) has less bias due to epistasis across inbreeding levels. However, the limitations to estimate epistasis are that we often obtain large negative values. Also, the correlations between first-and second-order statistics ((1/4)$\hat{\sigma}_A^2$ to (1/16)$\hat{\sigma}_{AA}^2$) might not be reliable and should be taken into account.

4.5 Design II

The design II mating design or factorial design was described by Comstock and Robinson (1948). It is also known as North Carolina design II or AB design. The assumptions for this mating design are similar to design I: no maternal effects,
linkage equilibrium, no epistasis, and arbitrary allele frequencies except when calculating level of dominance. However, this design has greater precision, it is more applicable to self-pollinated crops, and has a direct estimate of the level of dominance. On the other hand, it is more difficult to apply in open-pollinated species with only one inflorescence. Therefore, design II has not been used extensively for non-inbred plants in maize since multiple crosses on female plants are not possible. As a consequence, selfing of unselected S₀ plants is needed.

Hence, a representative sample of unselected S₁ progenies developed from a population can be used as females in crosses with unselected S₀ plants used as males (e.g. utilizing the same tassel several times). Seed from different S₁ plants is bulked for testing.

Males and females are taken at random if model II is followed and we make all possible crosses among different individuals. A diagram with four males and four females is used as an example. This can be one of many sets.

<table>
<thead>
<tr>
<th></th>
<th>m₁</th>
<th>m₂</th>
<th>m₃</th>
<th>m₄</th>
<th>Margins</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₁</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FULL – SIBS</td>
</tr>
<tr>
<td>f₃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HALF – SIBS</td>
</tr>
<tr>
<td>f₄</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HALF – SIBS</td>
</tr>
</tbody>
</table>

The crosses are among different individuals. If crosses are among the same individuals the design would be called diallel (explained later in the chapter). Basic features of the design II and diallel mating designs are quite different, but the genetic
### Table 4.6 Comparison of the diallel and design II mating designs for the possible crosses among parents

<table>
<thead>
<tr>
<th></th>
<th><strong>Diallel</strong></th>
<th></th>
<th><strong>Design II</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parents (males)</td>
<td></td>
<td>Parents (males)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Parents (females)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>—</td>
<td>X₁₂</td>
<td>X₁₃</td>
</tr>
<tr>
<td>2</td>
<td>X₂₁</td>
<td>—</td>
<td>X₂₃</td>
</tr>
<tr>
<td>3</td>
<td>X₃₁</td>
<td>X₃₂</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>X₄₁</td>
<td>X₄₂</td>
<td>X₄₃</td>
</tr>
</tbody>
</table>

Information obtained from the two designs is similar. For the diallel design, the same parents are used as males and females, whereas different sets of parents are used as males and females (see Table 4.6) for design II. If four parents are included in the diallel design, we have 6 cross-combinations and 12 cross-permutations. For comparison, if we include a set of eight parents in design II, we have 16 crosses (vs. 12 for all cross-permutations in the diallel design) but twice as many parents are included. With either design the number of crosses increases rapidly as the parents included increase, but the number of crosses is considerably less for design II, particularly when greater numbers of parents are used (Table 4.7). Approximately half as many crosses are produced when 10 or more parents are used. For a fixed number of experimental units, therefore, approximately twice as many parents can be used in the experiment. This is an advantage of design II, particularly if one wishes to estimate the genetic parameters of a reference population.

From Table 4.6 it is obvious we have a cross-classification design for analysis. Consequently, we will have sources of variation for males, females, and the interaction of males with females. Therefore, a factorial design is used to obtain expected mean squares in the ANOVA. The form of the analysis of variance when $m$ males are crossed with $f$ females and evaluated in $r$ replications is shown in Table 4.8.

The expected mean squares expressed in terms of the covariance of relatives are similar to those for the diallel analysis. The expectations of males and females for design II are equivalent to GCA, and the male $\times$ female source is equivalent to SCA of the diallel analysis. Because we have two sets of parents in design II, we have two independent estimates of GCA. Appropriate $F$-tests can be made to test for the differences among males and among females and for the interactions of males and females. Similar to the diallel analysis, the model I analysis provides estimates of GCA effects for males and females and SCA effects for males $\times$ females.

The model II analysis gives estimates of components of genetic variance that are estimable from covariances of relatives. From Table 4.8, $\hat{\sigma}^2_{Am} = \hat{\sigma}^2_{Af} = \hat{\text{Cov}}$ HS = ($\ell_4$$\hat{\sigma}^2_A$ for $F = 0$ and ($\ell_2$$\hat{\sigma}^2_A$ for $F = 1$, and $\hat{\sigma}^2_{mf} = \hat{\text{Cov}}$ FS $- \hat{\text{Cov}}$ HS$m - \hat{\text{Cov}}$ HS$f = ($\ell_4$$\hat{\sigma}^2_D$ for $F = 0$ and $\hat{\sigma}^2_D$ for $F = 1$; estimates are under the assumption of no epistasis in all instances. Two independent estimates of $\hat{\sigma}^2_A$ are calculated for $F = 0$ as $\hat{\sigma}^2_{Am} = 4(M_5 - M_3)/(r f)$ and $\hat{\sigma}^2_{Af} = 4(M_4 - M_3)/(r m)$. In summary,
Table 4.7  Number of crosses possible from the diallel and design II mating designs and possible options for different sample sizes of parents for partitioning into sets

<table>
<thead>
<tr>
<th>Number of parents</th>
<th>Sets of diallel</th>
<th>Sets of design II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>190</td>
<td>90</td>
</tr>
<tr>
<td>40</td>
<td>780</td>
<td>380</td>
</tr>
<tr>
<td>80</td>
<td>3,160</td>
<td>1,560</td>
</tr>
<tr>
<td>100</td>
<td>4,950</td>
<td>2,450</td>
</tr>
<tr>
<td>200</td>
<td>19,900</td>
<td>9,900</td>
</tr>
<tr>
<td>(n)</td>
<td>(n(n-1)/2)</td>
<td>(2[n'(n'-1)/2])</td>
</tr>
</tbody>
</table>

\(n'\) depends on the grouping of parents in sets

\(n'\) depends on the grouping of parents in sets

\(n\) Number of males and females are not equal
Table 4.8 Analysis of variance of the design II mating design in one environment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>Variance component</th>
<th>Covariance of relatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r-1</td>
<td>M_5</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}<em>{im}^2 + rf\hat{\sigma}</em>{m}^2 )</td>
<td>( \hat{\sigma}^2 + r[\hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HSf} - \hat{\text{Cov}} \text{ HSm}] + rf \hat{\text{Cov}} \text{ HSm} )</td>
</tr>
<tr>
<td>Males (M)</td>
<td>m-1</td>
<td>M_4</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}<em>{im}^2 + rm\hat{\sigma}</em>{f}^2 )</td>
<td>( \hat{\sigma}^2 + r[\hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HSf} - \hat{\text{Cov}} \text{ HSm}] + rm \hat{\text{Cov}} \text{ HSm} )</td>
</tr>
<tr>
<td>Females (F)</td>
<td>f-1</td>
<td>M_3</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}_{im}^2 )</td>
<td></td>
</tr>
<tr>
<td>M x F</td>
<td>(m-1)(f-1)</td>
<td>M_2</td>
<td>( \hat{\sigma}^2 )</td>
<td>( \hat{\sigma}^2 )</td>
</tr>
<tr>
<td>Error</td>
<td>(r-1)(mf-1)</td>
<td>M_1a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>rmf-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within plot</td>
<td>rmf(k-1)</td>
<td>M_1a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^aM_1\) is the within-plot mean square and includes the within-plot genetic variance (\( \hat{\sigma}_{wg}^2 \)) and environmental variance (\( \hat{\sigma}_{uw}^2 \)): \( \hat{\sigma}_{wg}^2 = \hat{\sigma}_G^2 - \hat{\text{Cov}} \text{ FS} \), thus \( \hat{\sigma}_G^2 = [\hat{\sigma}_{uw}^2 + (\hat{\sigma}_{im}^2 + \hat{\sigma}_{f}^2)] / k + \hat{\sigma}_p^2 \), where \( \hat{\sigma}_G \) is the total genetic variance, \( \hat{\sigma}_p^2 \) is the plot error variance, and \( k \) is the number of plants measured in each plot.

For \( F = 0 \)

Since \( \hat{\sigma}_m^2 \) and \( \hat{\sigma}_f^2 = \hat{\text{Cov}} \text{ HS} \) and \( \hat{\text{Cov}} \text{ HS} = (\frac{1}{4})\hat{\sigma}_A^2 \), \( \hat{\sigma}_A^2 = 4\hat{\sigma}_m^2 = 4\hat{\sigma}_f^2 \)

and,

\[ \hat{\sigma}_{im}^2 = \left[ \hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HSf} - \hat{\text{Cov}} \text{ HSm} \right] = \left( \frac{1}{4} \right) \hat{\sigma}_A^2 + \left( \frac{1}{4} \right) \hat{\sigma}_D^2 - \left( \frac{1}{4} \right) \hat{\sigma}_A^2 \]

Hence, we have a direct estimate of \( \hat{\sigma}_D^2 \)

\[ \hat{\sigma}_{im}^2 = \left( \frac{1}{4} \right) \hat{\sigma}_D^2, \quad \hat{\sigma}_D^2 = 4\hat{\sigma}_{im}^2 \]

The variance \( V \) of the estimates of \( \hat{\sigma}_A^2 \) are as follows:

\[ V(\hat{\sigma}_A^2) = \frac{16}{(rf)^2} \cdot 2 \left[ \frac{M_5^2}{m+1} + \frac{M_3^2}{(m-1)(f-1)+2} \right] \]

\[ V(\hat{\sigma}_A^2) = \frac{16}{(rm)^2} \cdot 2 \left[ \frac{M_4^2}{f+1} + \frac{M_3^2}{(m-1)(f-1)+2} \right] \]

When inbreeding of the parents is zero, \( \hat{\sigma}_D^2 \) is estimable as \( \hat{\sigma}_D^2 = 4(M_3 - M_2) / r \), which has a variance of

\[ V(\hat{\sigma}_D^2) = \frac{16}{r^2} \cdot 2 \left[ \frac{M_3^2}{(m-1)(f-1)+2} + \frac{M_2^2}{(r-1)(mf-1)+2} \right] \]
If epistasis is present then for $F = 0$

\[
\tilde{\text{Cov}}_{\text{HS}} = (\hat{ y}_4) \hat{\sigma}_A^2 + (\hat{ y}_6) \hat{\sigma}_{AA}^2, \quad \tilde{\text{Cov}}_{\text{FS}} = (\hat{ y}_2) \hat{\sigma}_D^2 + (\hat{ y}_4) \hat{\sigma}_{AA}^2 + (\hat{ y}_6) \hat{\sigma}_{AD}^2 + (\hat{ y}_8) \hat{\sigma}_{DD}^2
\]
\[
\hat{\sigma}_m^2 = \tilde{\text{Cov}}_{\text{HS}}
\]
\[
\hat{\sigma}_f^2 = \tilde{\text{Cov}}_{\text{HS}}
\]
\[
\hat{\sigma}_{fm}^2 = \tilde{\text{Cov}}_{\text{FS}} - \tilde{\text{Cov}}_{\text{HSf}} - \tilde{\text{Cov}}_{\text{HSm}} = (\hat{ y}_2) \hat{\sigma}_A^2 + (\hat{ y}_4) \hat{\sigma}_D^2 + (\hat{ y}_4) \hat{\sigma}_{AA}^2 + (\hat{ y}_6) \hat{\sigma}_{AD}^2 + (\hat{ y}_8) \hat{\sigma}_{DD}^2 - \frac{1}{2} [(\hat{ y}_2) \hat{\sigma}_A^2 + (\hat{ y}_6) \hat{\sigma}_{AA}^2] + \frac{(1/2) \hat{\sigma}_A^2}{\hat{\sigma}_D^2} + \frac{(1/4) \hat{\sigma}_A^2}{\hat{\sigma}_{DD}^2} + \frac{(1/8) \hat{\sigma}_A^2}{\hat{\sigma}_{AD}^2} + \frac{(1/16) \hat{\sigma}_A^2}{\hat{\sigma}_{DD}^2}
\]

and for $F = 1$

\[
\tilde{\text{Cov}}_{\text{HS}} = (\hat{ y}_2) \hat{\sigma}_A^2 + (\hat{ y}_4) \hat{\sigma}_{AA}^2
\]
\[
\tilde{\text{Cov}}_{\text{FS}} = \hat{\sigma}_D^2 + \hat{\sigma}_A^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{DD}^2
\]
\[
\hat{\sigma}_m^2 = \tilde{\text{Cov}}_{\text{HS}}
\]
\[
\hat{\sigma}_f^2 = \tilde{\text{Cov}}_{\text{HS}}
\]
\[
\hat{\sigma}_{fm}^2 = \tilde{\text{Cov}}_{\text{FS}} - \tilde{\text{Cov}}_{\text{HSf}} - \tilde{\text{Cov}}_{\text{HSm}} = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{DD}^2 - \left[ (\hat{ y}_2) \hat{\sigma}_A^2 + (\hat{ y}_6) \hat{\sigma}_{AA}^2 \right] + \left[ (\hat{ y}_2) \hat{\sigma}_A^2 \right] + \left[ \hat{\sigma}_D^2 \right] + \left[ \hat{\sigma}_{AD}^2 \right] + \left[ \hat{\sigma}_{DD}^2 \right]
\]

Estimates of components of variance for design II as well as other mating designs characterize the population from which the parents were a random sample. However, even though design II seems to merit further consideration, it has not been used nearly as extensively in maize as the diallel. The mechanics of making crosses when the parents are inbred lines are no different from those for diallel mating designs. For non-inbred S0 plants, however, multiple crosses on female plants are not possible in maize. As mentioned before, unselected S1 progenies developed from a population can be used in crosses with unselected S0 plants used as males; this requires making crosses on several plants (5–10) and bulking the seed for testing. Multiple pollinations from the S0 male plants are possible in maize, but care must be taken not to break or damage the male inflorescence if multiple pollinations are needed for more than 1 day. The main problem of crossing S0 plants as males onto a series of S1 progenies is coordinating the time of flowering because of protandry of males and delayed flowering of S1 progenies. It is suggested that the S0 plants be delay-planted to increase opportunities of simultaneous flowering of S1 progenies and S0 plants. Hallauer (1970) effectively used design II for estimation of genetic components of variance in maize populations by use of non-inbred parents. Care must be taken, however, to minimize selection in producing S1 progenies, in choice of S0 plants used as males, and in sufficient sampling of genotypes within each of the S1 progenies. Yang et al. (2010) have used three sets of design II for estimation of GCA and SCA effects (model I) in maize hybrids by use of inbred parents.
Design II has the following advantages over diallel designs if one is interested in estimating components of variance of a reference population: (1) more parents can be included for a given level of resources, (2) two independent estimates of $\hat{\sigma}_A^2$ are available, (3) an estimate of $\hat{\sigma}_D^2$ is determined directly from the mean squares, and (4) a greater number of parents can be included by subdividing parents into sets. Advantages (1) and (4) are related and can be used to increase the sampling of the reference population. As seen with design I, grouping of parents into sets permits pooling the sums of squares over sets. We are interested in obtaining estimates of components of variance rather than comparisons of means. If interest is primarily in estimation of genetic components of variance of a reference population, sets of diallel crosses also could be pooled so that advantages (1) and (4) relative to the diallel are not great. Generally, however, the diallel mating design assumes all possible crosses among a set of parents. For example, if 20 parents of a reference population are considered, we will have 190 crosses for the diallel and 100 crosses for design II (Table 4.7). But if we subdivide the 20 parents into two sets ($n' = 10$) of design II crosses, we will need to make only 50 crosses. Similarly, if we subdivide 20 parents into four sets ($n' = 5$) of 5 parents for diallel crossing, we would have 40 diallel crosses. Sampling 200 parents from a population would not be unreasonable: 40 five-parent ($n' = 5$) diallels result in 400 crosses and 20 ten-parent ($n' = 10$) design IIs result in 500 crosses. Table 4.7 summarizes some of the options made possible by partitioning the parents into sets of $n'$ parents, remembering that each design II set includes twice as many parents as each diallel.

The analysis of variance (Table 4.9) for parents grouped in sets includes a source due to sets, but the expectations of the mean squares of males, females, and males × females are the same for the components of variance and the covariances of relatives. An analysis is conducted on each set, and sums of squares and degrees of freedom are pooled over sets.

Repeating the experiments over environments gives the analysis of variance shown in Table 4.10. For the analysis repeated over environments, direct $F$-tests can be made for all sources of variation except for males and females. Satterthwaite’s

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Mean squares (MS)</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sets</td>
<td>$s-1^a$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replications/sets</td>
<td>$s(r-1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males/sets</td>
<td>$s(m-1)$</td>
<td>$M_4$</td>
<td>$\hat{\sigma}^2 + r\hat{a}^2_{fm} + rf\hat{a}^2_m$</td>
</tr>
<tr>
<td>Females/sets</td>
<td>$s(-1)$</td>
<td>$M_3$</td>
<td>$\hat{\sigma}^2 + r\hat{a}^2_{fm} + rm\hat{a}^2_f$</td>
</tr>
<tr>
<td>Males × females/sets</td>
<td>$s(m-1)(f-1)$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2 + r\hat{a}^2_{fm}$</td>
</tr>
<tr>
<td>Pooled error</td>
<td>$s(r-1)(mf-1)$</td>
<td>$M_1$</td>
<td>$\hat{\sigma}^2b$</td>
</tr>
<tr>
<td>Total</td>
<td>$s(rmf-1)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a, r, m, and f refer to the number of sets, replications, males, and females, respectively.

If individual plant data are taken, $\hat{\sigma}^2$ will be equal to $[(\hat{\sigma}^2_G - \text{Cov FS}) + \hat{\sigma}^2_{ae}] / k + \hat{\sigma}_p^2$, where $k$ is the number of plants measured per plot and $\hat{\sigma}_p^2$ is the experimental plot error.
Table 4.10 Analysis of variance of design II repeated over environments for model II

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments (E)</td>
<td>(e−1^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sets (S)</td>
<td>(s−1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S × E</td>
<td>((e−1)(s−1))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replications/S/E</td>
<td>(es(r−1))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males/S</td>
<td>((m−1))</td>
<td>(M_7) (\hat{\sigma}^2 + r\hat{\sigma}<em>{\text{ime}}^2 + rf\hat{\sigma}</em>{\text{me}}^2 + re\hat{\sigma}<em>{\text{mf}}^2 + rf\hat{\sigma}</em>{\text{m}}^2)</td>
<td></td>
</tr>
<tr>
<td>Females/S</td>
<td>((f−1))</td>
<td>(M_6) (\hat{\sigma}^2 + r\hat{\sigma}<em>{\text{ime}}^2 + rm\hat{\sigma}</em>{\text{lo}}^2 + re\hat{\sigma}<em>{\text{mr}}^2 + rf\hat{\sigma}</em>{\text{f}}^2)</td>
<td></td>
</tr>
<tr>
<td>Males × females/S</td>
<td>((m−1)(f−1))</td>
<td>(M_5) (\hat{\sigma}^2 + r\hat{\sigma}<em>{\text{ime}}^2 + re\hat{\sigma}</em>{\text{mf}}^2)</td>
<td></td>
</tr>
<tr>
<td>Males/S × E</td>
<td>((m−1)(e−1))</td>
<td>(M_4) (\hat{\sigma}^2 + r\hat{\sigma}<em>{\text{ime}}^2 + rf\hat{\sigma}</em>{\text{me}}^2)</td>
<td></td>
</tr>
<tr>
<td>Females/S × E</td>
<td>((f−1)(e−1))</td>
<td>(M_3) (\hat{\sigma}^2 + r\hat{\sigma}<em>{\text{fop}}^2 + rm\hat{\sigma}</em>{\text{lo}}^2)</td>
<td></td>
</tr>
<tr>
<td>Males × females/S × E</td>
<td>((m−1)(f−1)(e−1))</td>
<td>(M_2) (\hat{\sigma}^2 + r\hat{\sigma}_{\text{ime}}^2)</td>
<td></td>
</tr>
<tr>
<td>Pooled error</td>
<td>(es(r−1)(mf−1))</td>
<td>(M_1) (\hat{\sigma}^2)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>(esrmf−1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) \(e, s, r, m, a, n d f\) refer to the number of environments, sets within an environment, replications, males, and females, respectively

(1946) approximate test procedure can be used to synthesize mean squares that have the same expected value except for the effect being tested; e.g., for females, \(M_6 + M_2\) can be tested with \(M_4 + M_3\) with the following degrees of freedom:

\[
\begin{align*}
n_1 &= \frac{(M_6 + M_2)^2}{[M_6^2/[s(f − 1)] + M_2^2/[s(m − 1)(f − 1)(e − 1)]]} \\
n_2 &= \frac{(M_4 + M_3)^2}{[M_4^2/[s(m − 1)(e − 1)] + M_3^2/[s(f − 1)(e − 1)]]}
\end{align*}
\]

Tables 4.9 and 4.10 assume random effects of parents and environments; hence, estimates of the components of genetic variance are interpreted relative to the reference population and how they interact with environments. A measure of the average dominance of genes in the expression of the trait analyzed can be determined by the components of genetic variance estimated from the expected mean squares. If we assume a population in linkage equilibrium at \(p = q = 0.5\) and no epistasis, the components of genetic variance are estimated as

\[
\begin{align*}
\hat{\sigma}_A^2 &= \langle \frac{1}{2} \rangle \sum_i a_i^2 \\
\hat{\sigma}_D^2 &= \langle \frac{1}{4} \rangle \sum_i d_i^2
\end{align*}
\]

when,

\[
F = 0, \quad \hat{\sigma}_m^2 = \hat{\sigma}_f^2 = \hat{\sigma}_{\text{HMS}}^2 = \langle \frac{1}{4} \rangle \hat{\sigma}_A^2 \quad \text{and} \quad \hat{\sigma}_{mf}^2 = \text{Cov FS} - \text{Cov HS}_m - \text{Cov HS}_f = \langle \frac{1}{4} \rangle \hat{\sigma}_D^2.
\]

The average dominance of genes affecting the trait can be determined as
\[ d = \left( 2\hat{\sigma}_m^2 / \hat{\sigma}_m^2 \right)^{1/2} = \left( 2\hat{\sigma}_m^2 / \hat{\sigma}_f^2 \right)^{1/2} \]

From the calculated ratio we can determine which of the following levels of dominance of genes were operative: 0 is no dominance, 0 to 1 is partial dominance, 1 is complete dominance, and a value that exceeds 1 is termed overdominance. Examination of the assumptions for calculation of average dominance of genes shows that bias may be important. If digenic epistasis is present, the estimate of \( d \) will be biased upward because \( \text{Cov FS} - \text{Cov HS} - \text{Cov HSf} \) shows that we have a contribution of \((1/8)\hat{\sigma}_{AA}^2 + (1/16)\hat{\sigma}_{AD}^2 + (1/4)\hat{\sigma}_{DD}^2\) in the estimate of \( \hat{\sigma}_m^2 \) as follows:

If \( F = 0 \),

\[
\hat{\sigma}_m^2 = \text{Cov HS} = (\hat{\sigma}_A^2 + (1/4)\hat{\sigma}_{AA}^2)
\]
\[
\hat{\sigma}_f^2 = \text{Cov HS} = (\hat{\sigma}_A^2 + (1/8)\hat{\sigma}_{AD}^2 + (1/4)\hat{\sigma}_{DD}^2)
\]
\[
\hat{\sigma}_{fm}^2 = \text{Cov FS} - \text{Cov HS} - \text{Cov HSf} = (\hat{\sigma}_A^2 + (1/4)\hat{\sigma}_{AA}^2 + (1/4)\hat{\sigma}_{AD}^2 + (1/16)\hat{\sigma}_{DD}^2)
\]

And if \( F = 1 \),

\[
\hat{\sigma}_m^2 = \text{Cov HS} = (\hat{\sigma}_A^2 + (1/4)\hat{\sigma}_{AA}^2)
\]
\[
\hat{\sigma}_f^2 = \text{Cov HS} = (\hat{\sigma}_A^2 + (1/8)\hat{\sigma}_{AD}^2 + (1/4)\hat{\sigma}_{DD}^2)
\]
\[
\hat{\sigma}_{fm}^2 = \text{Cov FS} - \text{Cov HSf} - \text{Cov HSf} = (\hat{\sigma}_A^2 + 2\hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \hat{\sigma}_{AD}^2 + \hat{\sigma}_{DD}^2)
\]

and compares positively with design I.

Effects of linkage bias depend to some extent on the reference population under consideration. If a large random mating population is considered, linkage bias probably is minimal; whereas in an F2 population created from two inbred lines, linkage disequilibrium may be important. Coupling phase linkages would not be a source of bias in estimates of \( d \) because both \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) have a positive or upward bias if the population is in linkage disequilibrium, i.e., \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) are biased but \( d \) is not. If the genes are in repulsion phase linkages (which is more likely when crossing two inbred lines to correct weaknesses in both), the expression could be the same as for overdominance in the expression of independently segregating genes, although none of the linked genes was individually more than partially dominant to their alleles. As seen before, repulsion phase linkages cause an upward or positive bias in the estimate of \( \hat{\sigma}_D^2 \) (same as coupling phase linkages) but cause a downward or negative bias in estimates of \( \hat{\sigma}_A^2 \). Hence, \( \hat{\sigma}_D^2 \) will be overestimated and \( \hat{\sigma}_A^2 \) underestimated, which
results in an overestimate of $d$. Gene frequencies will not be known but are expected
to be approximately $p = q = 0.5$ for F2 populations and may not diverge greatly from
0.5 for populations that have not been under intentional selection. Populations under
long-term selection pressures may have a significant departure from 0.5, however.

Estimates of heritability can be calculated from use of estimates of $\hat{\sigma}_A^2$ from male
and female components of variance. Assuming non-inbred parents and no epistasis,
$\hat{\sigma}_A^2 = 4\hat{\sigma}_m^2$ (Table 4.9) and an estimate of $\hat{h}^2$ based on the mean of $r$ plots (an
entry-mean basis) for one environment is

$$\hat{h}^2 = \frac{4\hat{\sigma}_m^2}{\hat{\sigma}^2/r + 4\hat{\sigma}_{mf}^2 + 4\hat{\sigma}_m^2}$$

which has an approximate standard error of

$$SE(\hat{h}^2) = \frac{4SE(\hat{\sigma}_m^2)}{\hat{\sigma}^2/r + 4\hat{\sigma}_{mf}^2 + 4\hat{\sigma}_m^2}$$

A similar estimate of heritability can be calculated from the female source of vari-
ation. Also, if the degrees of freedom are equal for male and female sources of
variation, an estimate of $\hat{\sigma}_A^2$ can be made from the mean square obtained by pooling
the male and female degrees of freedom and sums of squares. Standard errors of
components of variance are calculated in the usual manner, and the standard error
of $\hat{\sigma}_A^2$ obtained from the mean square by pooling the males and females would be
reduced because twice as many degrees of freedom are included in the denomina-
tor. If individual plant data are collected, an estimate of heritability on an individual
plant basis can be calculated as

$$\hat{h}^2 = \frac{4\hat{\sigma}_m^2}{\hat{\sigma}^2_r + \hat{\sigma}_p^2 + \hat{\sigma}_{mf}^2 + \hat{\sigma}_f^2 + \hat{\sigma}_m^2}$$

where $\hat{\sigma}_p^2$ is plot error
and $\hat{\sigma}_w^2 + \hat{\sigma}_{we}^2 + (\hat{\sigma}_G^2 - \hat{\text{Cov FS}})$. If the parents are non-inbred ($F = 0$), $\hat{\sigma}_w^2$
includes environmental variance among plants within a plot ($\hat{\sigma}_{we}^2$) and the genetic variance
within plots, ($\hat{\sigma}_G^2 - \hat{\text{Cov FS}}) = (\frac{N}{2})\hat{\sigma}_A^2 + (\frac{N}{2})\hat{\sigma}_D^2$, assuming no epistasis. If the par-
ents are homozygous, $\hat{\sigma}_w^2$ of course includes only $\hat{\sigma}_{we}^2$. The standard error of the
individual plant heritability estimate is

$$SE(\hat{h}^2) = 4SE(\hat{\sigma}_m^2)/(\hat{\sigma}_w^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{mf}^2 + \hat{\sigma}_f^2 + \hat{\sigma}_m^2).$$

Estimates of heritability unbiased by genotype–environment interaction and based
on the mean of $re$ plots can also be calculated from Table 4.10. An estimate
is available from either the male or female component of variance, but the best esti-
mate would be obtained from pooling male and female sums of squares. As Nyquist
(1991) and Holland et al. (2003) have emphasized, a more meaningful and useful
estimate of heritability for genetic improvement programs are those based on half-sib and full-sib family means. For half-sib family means for the male source of variation, the estimate of heritability would be

\[
\hat{h}^2 = \frac{\hat{\sigma}_m^2}{\hat{\sigma}_m^2/ef + \hat{\sigma}_m^2/e + \hat{\sigma}_m^2/f + \hat{\sigma}_m^2} \quad \text{Across environments (Table 4.10) Half-sib progeny means basis}
\]

A similar estimate of heritability can be calculated from the female source of variation except \( m \) would be the divisor in the denominator components of variance. If the male and female sources are pooled, as done to estimate the half-sib variance component, a better estimate of heritability based on half-sib family means could be calculated. For the male \( \times \) female source of variation, an estimate of heritability on full-sib family means can also be calculated as follows:

\[
\hat{h}^2 = \frac{\hat{\sigma}_{fm}^2}{\hat{\sigma}_{fm}^2/re + \hat{\sigma}_{efm}^2/r + \hat{\sigma}_{fm}^2} \quad \text{Across environments (Table 4.10) Full-sib progeny means basis}
\]

These estimates of heritability can be used to predict genetic response to selection based on either full-sib or half-sib family means (Nyquist, 1991); this is the primary reason for the estimation of the genetic components of variance (Chapter 6). Exact confidence limits for the estimates of heritability can be calculated as shown my Knapp et al. (1985).

Design II is very useful for estimation of genetic variances in a population. In plants we usually are able to have balanced data, and pooling the male and female sums of squares provides an estimate of \( \hat{\sigma}_A^2 \) with reasonable errors. If only a few selected parents are included, design II has no advantages over the diallel for estimating genetic effects of parents (GCA) and their crosses (SCA); the same information can be obtained from the diallel and design II. Design II also is useful when there is a grouping of male-sterile and male-fertile parents, such as the A and B lines of *Sorghum bicolor*.

### 4.6 Design III

The design III mating design also was developed by Comstock and Robinson (1948). This specific mating design was made with the purpose to estimate the average level of dominance of genes affecting a trait (Table 4.11). However, it also provides good estimates of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) for F2 populations assuming absence of linkage and epistasis. The advantage of design III over previous ones is that the estimation of dominance is not subjected to any assumption regarding allele frequencies.

Genes for F2 populations that are developed from crossing two inbred lines often were thought to be in the overdominant range when average level of dominance was estimated. Estimates of average level of dominance assumed linkage equilibrium of the populations. Because F2 populations were sampled, linkage effects could have
established a serious bias in estimating $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. If linkage effects are present, estimates of $\hat{\sigma}_D^2$ are always biased positively, regardless of whether linkage is in the coupling or the repulsion phase. Linkage bias for estimates of $\hat{\sigma}_A^2$, however, depends on the phase of linkage; $\hat{\sigma}_A^2$ is underestimated for repulsion phase linkages and overestimated for coupling phase linkages. Because the measure of dominance of genes was in the overdominant range it was speculated that repulsion phase linkages had overestimated $\hat{\sigma}_D^2$ and underestimated $\hat{\sigma}_A^2$. The design III mating design, therefore, has primarily been used in maize F2 populations to determine the effects of linkages on the estimates of $\hat{\sigma}_A^2$, $\hat{\sigma}_D^2$, and average level of dominance. The initial reference population for mating design III, therefore, is an F2 population. Backcrossing randomly chosen S0 plants (males) of the F2 population to both homozygous parents (P1 and P2 used to develop the F2 population, females) develops progenies for this design. Therefore, a random and large enough sample of F2 seeds is planted to make reciprocal backcrosses. As a consequence, pairs of progenies (one pair per F2 male) are the progenies utilized for evaluation.

![Diagram of mating designs](image)

### Table 4.11 Estimation of level of dominance across North Carolina mating designs

<table>
<thead>
<tr>
<th>Design I</th>
<th>Design II</th>
<th>Design III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = q = \frac{1}{2}$</td>
<td>$p = q = \frac{1}{2}$</td>
<td>No assumptions</td>
</tr>
<tr>
<td>$\bar{d} = \sqrt{\frac{2\hat{\sigma}_D^2}{\hat{\sigma}_m^2}}$</td>
<td>$\bar{d} = \sqrt{\frac{2\hat{\sigma}_{mf}^2}{\hat{\sigma}_m^2}}$ (or $\hat{\sigma}_f^2$)</td>
<td>$\bar{d} = \sqrt{\frac{2\hat{\sigma}_D^2}{\hat{\sigma}<em>m^2} \div 2\hat{\sigma}</em>{im}^2 - \hat{\sigma}_m^2}$</td>
</tr>
</tbody>
</table>
Therefore, X equals to \( m \) \( F_2 \) males and the number of parents used in the backcrosses is always equal to 2 and becomes a fixed factor. Entries include \( m \) pairs of progenies, one pair for each \( F_2 \) male, and the parental lines. The ANOVA for design III replicated in one environment is shown in Table 4.12

**Table 4.12** Analysis of variance of design III progenies tested in one environment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>Variance components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>( r-1^a )</td>
<td></td>
<td>( M_4 ) ( \hat{\sigma}^2 + r\hat{\sigma}^2_{mp} + rmK_p^2 )</td>
</tr>
<tr>
<td>Parents (p)</td>
<td>1</td>
<td>( M_3 )</td>
<td>( \hat{\sigma}^2 + 2r\hat{\sigma}^2_m )</td>
</tr>
<tr>
<td>Males (m)</td>
<td>( m-1 )</td>
<td>( M_2 )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_{mp} )</td>
</tr>
<tr>
<td>m × p</td>
<td>( m-1 )</td>
<td>( M_1 )</td>
<td>( \hat{\sigma}^2 )</td>
</tr>
<tr>
<td>Error</td>
<td>((r-1)(2m-1))</td>
<td>1</td>
<td>( \hat{\sigma}^2 )</td>
</tr>
<tr>
<td>Total</td>
<td>( 2mr-1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^a r \) and \( m \) refer to number of replications and male plants, respectively

Mean square expectations to focus on in design III are the component of variance among males (\( \hat{\sigma}^2_m \)) and the one for the interaction of males and inbred parents (\( \hat{\sigma}^2_{mp} \)). Direct \( F \)-tests are possible for \( \hat{\sigma}^2_{mp} \) and \( \hat{\sigma}^2_m \) with the error term but they do not give us knowledge of the genetic structure of the progenies and how they relate to components of variance. Comstock and Robinson (1952) defined the genetic structure of progenies in the absence of linkage (independence of segregation) and epistasis (independence of action) as

\[
\hat{\sigma}^2_m = \frac{1}{2} \sum pq a^2 \\
\hat{\sigma}^2_{mp} = \sum pq d^2
\]

Therefore, genetic parameters can be estimated as in other designs. Since \( p = q = \frac{1}{2} \) is expected for \( F_2 \) populations, then

\[
\hat{\sigma}^2_m [(M_3-M_1)/2r] = (1/8) \sum a^2 = (\frac{1}{4})\hat{\sigma}^2_A \\
\hat{\sigma}^2_{mp} [(M_2-M_1)/r] = (1/4) \sum d^2 = \hat{\sigma}^2_D \\
\hat{\sigma}^2_A = 4 \hat{\sigma}^2_m \\
\hat{\sigma}^2_D = \hat{\sigma}^2_{mp}
\]

Thus, \( \hat{\sigma}^2_A = 4\hat{\sigma}^2_m \) and \( \hat{\sigma}^2_{mp} = \hat{\sigma}^2_D \). From the expectations of the \( \hat{\sigma}^2_m \) and \( \hat{\sigma}^2_{mp} \) components of variance, a measure of the dominance of genes can be obtained and the interpretation of the ratio is the same as for design II.
Design III provides exact F-tests of two hypotheses concerning the relative importance of dominance effects: (1) that dominance is not present (this can be tested by comparison of the $M_1$ and $M_2$ mean squares of Table 4.12; except for sampling errors the $M_2$ mean square will be greater than the $M_1$ mean square only when some level of dominance occurs at one or more loci) and (2) that dominance is complete (which is sometimes assumed for genes determining quantitative traits that exhibit heterosis). For the assumption of independent segregation of loci and independent action of genes (no epistasis), expected values of $M_3$ and $M_2$ are $M_3 = \hat{\sigma}_A^2 + 2r\hat{\sigma}_m^2 = \hat{\sigma}_A^2 + (r/4)\sum_i a_i^2$ and $M_2 = \hat{\sigma}_A^2 + (r/4)\sum_i d_i^2$. Expected mean squares are similar when $\sum_i a_i^2 = \sum_i d_i^2$; and deviations of the ratio, $M_3/M_2$, from unity in either direction indicate that $\sum_i a_i^2 \neq \sum_i d_i^2$ and that dominance is not complete.

From Table 4.12 the average level of dominance can be estimated as $[(M_2 - M_1)/(M_3 - M_1)]^{1/2}$. When testing level of dominance for genes in F2 populations these were often in the overdominant range. However, testing of linkage showed that estimates of overdominance in these populations were not true (pseudo-overdominance). The joint action of linked genes was suggested instead of overdominance at individual loci. Therefore, linkage biases are important in F2 populations.

It is valid to estimate heritabilities of the traits measured in the design III analysis because the F2 population is a valid reference population. Sufficient sampling is needed to obtain valid estimates of components of variance to determine average level of dominance. To sample the F2 populations adequately, the number of pairs of progenies may be greater than we desire to include in one replication. The same procedures used for designs I and II for local control of experimental error can be used for design III by grouping pairs of progenies into sets. Each set, therefore, is analyzed as in Table 4.12 and then sums of squares and degrees of freedom across sets are pooled. Hence, the components are useful for estimating heritability in the narrow sense in the F2 population. From the components given in Table 4.12, an estimate of heritability based on the mean of $r$ plots can be determined as follows:

$$\hat{h}^2 = 4\hat{\sigma}_m^2 / (\hat{\sigma}_A^2 / r + \hat{\sigma}_m^2 + 4\hat!\sigma_m^2).$$

If the experiment is repeated across environments, the analysis is as shown in Table 4.13. Combined analyses across environments provide estimates of the interaction of the additive and dominance effects with environments as seen in design I and II. Direct F-tests are available for each mean square and components of variance can be calculated directly from the mean squares with their appropriate standard errors. In addition to providing a measure of the dominance of genes for the expression of a trait, design III also is an excellent mating design for estimation of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ for an F2 population, assuming no linkage and epistasis. The combined analysis provides estimates of the interaction of the additive and dominance effects with environments (Table 4.13).

Variance components of design III are useful for estimating heritability in the narrow sense in the F2 population. Estimates of heritabilities in which the estimate of $\hat{\sigma}_A^2$ is not biased by $\hat{\sigma}_{AE}^2$ can be obtained if the experiment is repeated over environments (Table 4.13) as
4.6 Design III

Table 4.13 Analysis of variance of design III progenies pooled over sets within environments and combined across environments

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments (E)</td>
<td>e−1(^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sets</td>
<td>s−1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E × sets</td>
<td>(e−1)(s−1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replications/sets/E</td>
<td>se(r−1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents/sets</td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E × parents/sets</td>
<td>s(e−1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males/sets</td>
<td>s(m−1)</td>
<td></td>
<td>(\hat{\sigma}^2 + 2\hat{\sigma}<em>{me}^2 + 2re\hat{\sigma}</em>{m}^2)</td>
</tr>
<tr>
<td>Males × parents/sets</td>
<td>s(m−1)</td>
<td>(M_5)</td>
<td>(\hat{\sigma}^2 + 2r\hat{\sigma}<em>{me}^2 + re\hat{\sigma}</em>{mp}^2)</td>
</tr>
<tr>
<td>E × males/sets</td>
<td>(e−1)s(m−1)</td>
<td>(M_4)</td>
<td>(\hat{\sigma}^2 + 2r\hat{\sigma}_{me}^2)</td>
</tr>
<tr>
<td>E × males × parents/sets</td>
<td>(e−1)(s−1)</td>
<td>(M_3)</td>
<td>(\hat{\sigma}^2 + 2r\hat{\sigma}_{mp}^2)</td>
</tr>
<tr>
<td>Pooled error</td>
<td>es(r−1)(2m−1)</td>
<td>(M_2)</td>
<td>(\hat{\sigma}^2)</td>
</tr>
<tr>
<td>Total</td>
<td>es2mr−1</td>
<td></td>
<td>(\hat{\sigma}^2)</td>
</tr>
</tbody>
</table>

\(^a\) e, s, r, and m refer to the number of environments, sets within an environment, replications, and pairs of progenies for each male parent, respectively.

\[\hat{h}^2 = \frac{4\hat{\sigma}_{m}^2}{\hat{\sigma}^2 / (re) + \hat{\sigma}_{mpe}^2 / e + \hat{\sigma}_{mp}^2 / e + 4\hat{\sigma}_{me}^2 / e + 4\hat{\sigma}_{m}^2}\]

which is on the mean basis of re plots. Approximate standard errors of heritability estimates can be obtained as previously shown for the other mating designs.

The primary goal of the design III is to estimate relative dominance of alleles and to simplify the assumption of allele frequencies of \(p = q = 0.5\) for segregating loci in crosses of inbred lines to form the F\(_2\) populations. This simplification may not seem appropriate in some instances, but greater emphasis in applied maize breeding programs is given to selection within F\(_2\) populations developed from elite-line crosses within identified heterotic groups. Hence, estimates of heritability for different traits within F\(_2\) populations are of importance during inbreeding and selection for developing recycled lines (e.g., Fig. 1.4). Individual F\(_2\) plants as males are crossed to both inbred parents, and the differences among male means are covariances of half-sib families. Similar to the other mating designs, heritability estimates based on half-sib family means are

\[\hat{h}^2 = \frac{\hat{\sigma}_{m}^2}{\hat{\sigma}^2 / r + \hat{\sigma}_{m}^2}\] Within environments (Table 4.12)

\[\hat{h}^2 = \frac{\hat{\sigma}_{m}^2}{\hat{\sigma}^2 / 2re + \hat{\sigma}_{mpe}^2 / 2e + \hat{\sigma}_{mp}^2 / e + \hat{\sigma}_{me}^2 / e + \hat{\sigma}_{m}^2}\] Across environments (Table 4.13)

The estimates would be specific for each F\(_2\) population, but a summary of the relative heritability estimates for F\(_2\) populations would provide general guidelines for
the relative expected effectiveness of selection among progenies during inbreeding (e.g., Bauman, 1981).

Design III is powerful in testing for presence of dominance effects, but linkage biases may be serious in the estimation of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ in the F2 populations where effects of linkage are expected to be at a maximum. Tests for effects of linkage biases can be made by advancing the F2 and successive generations by random mating (Gardner et al., 1953; Gardner, 1963). Random mating can be done by either isolation plantings or hand-pollination but with no selection of male and female plants in the matings. If the first random mating in the F2 population is designated as synthetic 1 (Syn 1), eight generations of random mating of plants (Syn 8), for example, permit recombination of linked loci and, at least for loose linkages, linkage equilibrium would be approached. The approach to linkage equilibrium obviously depends on the rate of recombination and may require many generations of random mating for tightly linked gene combinations. After the F2 population has been advanced by random mating for, say, eight generations, the design III mating scheme is repeated for the F2 and F2 Syn 8 population. Estimates of $\hat{\sigma}_A^2$, $\hat{\sigma}_D^2$, and $\hat{d}$ from the analysis of the two populations are compared to determine effects of linkage bias on the estimates. Gardner and Lonnquist (1959) have reported estimates testing effects of linkage and found that they were an important bias in the estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. Hence estimates of overdominance in the F2 populations were pseudo-overdominance because of the joint action of linked genes rather than overdominance at individual loci. As an example, say 200 F2 individuals are intermated to test the importance of dominance and effects of linkage bias. Hand-pollination without selection is performed among individuals to develop Synthetic 1 (first random mating generation or Syn 1). Five generations of random mating allow for recombination of most linked loci since the rate of recombination in this population is high. Design III is performed for the F2 Syn 5 and is repeated for the original F2 population. The following will show the type of situations to expect:

(1) Data show that the average level of dominance for the F2 generation and the one for the F2 Syn 5 is higher than 0 but lower than 1. The difference, though, is non-significant.

In this case, we can expect partial dominance to be important (no overdominance).

(2) The difference is again non-significant but the values are higher than 0.5 and lower than 1.

Here, we can expect the majority of loci having partial dominance and a few loci may have overdominance effects.

(3) The difference is significant. Level of dominance for the F2 generation is greater than 1 and level of dominance for the F2 Syn 5 is less than 1.

Some overdominance effects are possible and linkage disequilibrium would also have impact.
4.7 Diallel Methods

The diallel mating design has been used and abused more extensively than any other in maize and other plant species. However, it can be very useful if properly analyzed and interpreted. Although extensive theoretical research and discussion have been presented, the main problem seems to arise from interpretations and inferences that can be made about estimates obtained from analysis of the diallel crosses.

As the name implies, crosses are made in pairs for \( n \) number of parents in a similar way done for design II. However, the main difference with design II is that the same individuals are used as parents for the diallel mating design (individuals are used both as male and female, see below).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Margins Yj</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>( Y_1 )</td>
</tr>
<tr>
<td>2</td>
<td>2.1</td>
<td>2.2</td>
<td>2.3</td>
<td>2.4</td>
<td>( Y_2 )</td>
</tr>
<tr>
<td>3</td>
<td>3.1</td>
<td>3.2</td>
<td>3.3</td>
<td>3.4</td>
<td>( Y_3 )</td>
</tr>
<tr>
<td>4</td>
<td>4.1</td>
<td>4.2</td>
<td>4.3</td>
<td>4.4</td>
<td>( Y_4 )</td>
</tr>
<tr>
<td>Margins Yj</td>
<td>( Y_1 )</td>
<td>( Y_2 )</td>
<td>( Y_3 )</td>
<td>( Y_4 )</td>
<td>( Y )</td>
</tr>
</tbody>
</table>

Diallel crossing schemes and analyses have been developed for parents that range from inbred lines to broad genetic base varieties. After the crosses are made, evaluated, and analyzed, inferences regarding the types of gene action can be made. It is important, however, that the assumptions and limitations of the diallel mating design are realized when one interprets the data. If correctly analyzed, the diallel mating design is very powerful, e.g., alternative heterotic patterns have been proposed (Hallauer et al., 1988; Carena, 2005; Melani and Carena, 2005; Carena and Wicks III, 2006).

First, let us consider how the crosses are made for the diallel mating design. If one has a complete diallel, it will include all possible crosses and parents. If \( n \) parents are included in the diallel and crossed in pairs, the number of permutations of \( n \) parents taken two at a time becomes \( n!(n-1)! \) or \( n!/(n-2)! \) because \( r = 2 \), which reduces to \( n(n-1) \). For this case the reciprocal crosses (e.g., 1 \( \times \) 2 and 2 \( \times \) 1 for parents 1 and 2) are made among the \( n \) parents. If reciprocal crosses are not included in the crosses among \( n \) parents, the possible arrangements of \( n \) parents taken \( r \) at a time become \( n!/[r!(n-r)!!] \) or \( n!/2 \) because \( r = 2 \); this reduces to \( n(n-1)/2 \). Assume 10 parents are included for diallel crossing; we would have 90 crosses if reciprocals are made and 45 if reciprocals are ignored. If the 10 parents are included for evaluation we would have 100 entries for the first case and 55 for the second.
Table 4.14 Possible number of entries from the diallel mating design for different numbers of parents

<table>
<thead>
<tr>
<th>Number of parents</th>
<th>Crosses</th>
<th>Total&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Permutations</th>
<th>Crosses</th>
<th>Total&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>21</td>
<td>30</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>28</td>
<td>42</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>37</td>
<td>56</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>45</td>
<td>72</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>55</td>
<td>90</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>105</td>
<td>120</td>
<td>210</td>
<td>225</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>190</td>
<td>210</td>
<td>380</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1225</td>
<td>1275</td>
<td>2450</td>
<td>2500</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>4950</td>
<td>5050</td>
<td>9900</td>
<td>10,000</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>n(n−1)/2</td>
<td>n(n+1)/2</td>
<td>n(n−1)</td>
<td>n&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Total number of entries refers to all cross-combinations and parents, i.e., \( n + n(n−1)/2 = n(n+1)/2 \)

<sup>b</sup>Total number of entries refers to all cross-permutations and parents, i.e., \( n + n(n−1) = n^2 \)

Table 4.14 summarizes several possible examples of the entries available from the diallel crossing of different numbers of parents.

An alternative way of looking at the diallel mating design is through different options or methods (Griffing, 1956). Following the example of \( n = 10 \) different types of progenies can be produced with the diallel mating design. As a consequence, different analyses can be used. There are four methods of producing progenies:

a) Method I = \( n^2 \) (10<sup>2</sup> = 100). It includes all possible crosses and parents.

b) Method II = \( n(n+1)/2 \) (10 × \( \frac{1}{2} \times 11 = 55 \)). This method is the most widely used and it includes one set of crosses and the parents (no reciprocals).

c) Method III = \( n(n−1) \) (10 × 9 = 90). It includes two sets of crosses without parents.

d) Method IV = \( n(n−1)/2 \) (10 × \( \frac{1}{2} \times 45 = 45 \)). It only includes one set of crosses with neither reciprocals nor parents.

The option will change depending on the material used. In maize, for pure lines the most logical choice would be to use one or two sets of crosses without parents. Otherwise, competition effects would be important. Contrarily, if we use synthetic varieties we can use diallel mating designs including not only crosses but also parents to compare mean performance and heterosis. Other species with less inbreeding depression might take advantage of using designs including parents and crosses when \( F = 1 \) (e.g., sorghum).

Based on the previous information we can see that one limitation of the diallel design is the number of parents that can practically be included (Table 4.15).
As the number of parents increases the number of possible crosses increases significantly. Therefore, with the diallel mating design it becomes nearly impossible to estimate the genetic variation present within a population. Obviously, as the number of parents increases, the number of possible crosses increases very rapidly. After one includes 10 to 15 parents, the number of crosses to make and to evaluate becomes somewhat unmanageable. Should one want to estimate the genetic variation present within a population, it would not be unreasonable to include 100 individuals to represent the range of genotypes within the population. The inclusion of only the combination of crosses (ignoring reciprocals) among 100 parents would require making and evaluating 4950 crosses for a complete diallel. It is readily apparent that the number of parents one can include to produce the crosses is an important factor in the diallel mating design.

Sample size is important when making crosses of non-inbred parents. In maize, approximately 100–200 plants are needed to adequately sample the genotypes in a population. Diallel designs among inbred lines are similar across species. However, the degree of difficulty in making the crosses and the amount of seed produced might be limiting factors to consider. To increase the number of seed in self-pollinated species one might choose to self $F_1$ crosses but the progeny will have $F = 0.5$ having this an impact on covariance of relatives.

The mechanical procedures for making the diallel crosses will vary among crop species (self- vs. cross-pollinators) and within crop species (inbred vs. non-inbred parents). If the parents are relatively homozygous (inbred lines), the series of diallel crosses can be made by repeating each parent for each combination of crosses and making paired-row crosses; the only limitation to the number of plants included and cross-pollinated for each pair-row cross is the quantity of seed needed for testing the crosses. By use of paired-row crosses, seed produced on each parent can be bulked for each cross-combination or kept separate if each cross-permutation is desired. Diallel crosses among a set of maize inbred lines are usually not too difficult, provided the proper timing is made for the flowering of the male and female inflorescences.

### Table 4.15  Potential number of parents across diallel mating design methods

<table>
<thead>
<tr>
<th>Number of parents</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>225</td>
</tr>
<tr>
<td>20</td>
<td>400</td>
</tr>
<tr>
<td>50</td>
<td>2500</td>
</tr>
<tr>
<td>100</td>
<td>10000</td>
</tr>
<tr>
<td>$n$</td>
<td>$n^2$</td>
</tr>
</tbody>
</table>

As the number of parents increases the number of possible crosses increases significantly. Therefore, with the diallel mating design it becomes nearly impossible to estimate the genetic variation present within a population. Obviously, as the number of parents increases, the number of possible crosses increases very rapidly. After one includes 10 to 15 parents, the number of crosses to make and to evaluate becomes somewhat unmanageable. Should one want to estimate the genetic variation present within a population, it would not be unreasonable to include 100 individuals to represent the range of genotypes within the population. The inclusion of only the combination of crosses (ignoring reciprocals) among 100 parents would require making and evaluating 4950 crosses for a complete diallel. It is readily apparent that the number of parents one can include to produce the crosses is an important factor in the diallel mating design.

Sample size is important when making crosses of non-inbred parents. In maize, approximately 100–200 plants are needed to adequately sample the genotypes in a population. Diallel designs among inbred lines are similar across species. However, the degree of difficulty in making the crosses and the amount of seed produced might be limiting factors to consider. To increase the number of seed in self-pollinated species one might choose to self $F_1$ crosses but the progeny will have $F = 0.5$ having this an impact on covariance of relatives.

The mechanical procedures for making the diallel crosses will vary among crop species (self- vs. cross-pollinators) and within crop species (inbred vs. non-inbred parents). If the parents are relatively homozygous (inbred lines), the series of diallel crosses can be made by repeating each parent for each combination of crosses and making paired-row crosses; the only limitation to the number of plants included and cross-pollinated for each pair-row cross is the quantity of seed needed for testing the crosses. By use of paired-row crosses, seed produced on each parent can be bulked for each cross-combination or kept separate if each cross-permutation is desired. Diallel crosses among a set of maize inbred lines are usually not too difficult, provided the proper timing is made for the flowering of the male and female inflorescences.
Diallel crosses among a set of maize populations are handled similarly to inbred lines, but the sampling of the population genotypes increases the number of individual plants included in the population crosses. Amount of seed usually is not a problem, but the number of crosses between different plants required to sample the populations increases the space and time needed. Several sets of pair-rows per cross are recommended to increase the sample size. Also, detasseling males after crossing can make the sample more representative with the advantage of reducing future number of pollinations. Shootbags from males can also be removed. Crosses between 10 plants of inbred lines may be sufficient for seed needs, whereas many more are necessary to adequately sample the genotypes in a population. Diallel crosses among pure lines of *Avena, Hordeum, Triticum,* and *Glycine,* for example, would be similar to inbred lines of maize, but the degree of difficulty in making the crosses often is the limiting factor. For a multiflowered self-pollinated species such as *Nicotiana,* crosses can be made quite easily with sufficient quantities of seed and small numbers of plants. It is obvious, therefore, that the diallel mating design can be used for most crop species, and the extent of its use depends on the difficulty of crossing and the resultant seed supplies. Seed produced on each set of crosses between parents can be either kept separate if one desires to test for reciprocal and maternal effects or bulked if one is interested only in performance of the crosses.

Genotypes produced by any of the methods of the diallel mating design are grown in replicated tests based on randomized complete or incomplete block designs. The diallel series of crosses are grown in replicated tests, with appropriate randomizations, to determine the relative merit of the parents in crosses. If 10 or less parents are included in the diallel crosses, a randomized complete block design should be satisfactory in most instances where uniform land is present. Incomplete block designs should be considered if the number of crosses and the environmental variability among experimental units are large. If we assume that only crosses among parents are tested in $e$ environments, the initial analysis of variance to determine if the variation among crosses is significantly different from zero is shown in Table 4.16. If no significant differences exist, there is no need to proceed further because apparently no detectable differences were contributed by the parents to their offspring.

### Table 4.16  Analysis of variance of a diallel set of crosses among $n$ parents evaluated in $e$ environments (method IV)

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>E(MS)</th>
<th>Model I</th>
<th>Model II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments (E)</td>
<td>$e−1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replications/E</td>
<td>$e(r−1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crosses (C)</td>
<td>$[n(n−1)/2]−1$</td>
<td>$M_3$</td>
<td>$\hat{\sigma}^2 + r\hat{K}_{C}^2$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_{C}^2 + re\hat{\sigma}^2$</td>
</tr>
<tr>
<td>C × E</td>
<td>$(e−1)[(n(n−1)/2]−1]$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}^2 + r\hat{K}_{CE}^2$</td>
<td>$\hat{\sigma}^2 + r\hat{\sigma}_{CE}^2$</td>
</tr>
<tr>
<td>Pooled error</td>
<td>$e(r−1)[(n(n−1)/2]−1$</td>
<td>$M_1$</td>
<td>$\hat{\sigma}^2$</td>
<td>$\hat{\sigma}^2$</td>
</tr>
<tr>
<td>Total</td>
<td>$er[n(n−1)/2]−1$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Before the experiments were conducted, an important decision was made about the parents included to make the crosses: Are the parents the reference genotypes or are the parents random genotypes from some reference population? Parents can be either the reference genotypes (model I or fixed model) or random genotypes from a reference population (model II or random model). This decision is made before analysis and the interpretation of the analysis changes depending on that decision. The answer to the former question has great implications in the interpretations made from the analysis of the diallel mating design, and it usually has been the basic feature in arguments for and against the utility of that design to provide the information desired by the researcher. Usually, the assumption made about the parents to be included, not how the experiment was conducted and analyzed, causes difficulties in the interpretation of the estimated parameters. Griffing (1956) and Cockerham (1963) have discussed the diallel analysis in detail as well as the analysis of variance for fixed models (model I, where the parents are the genotypes under consideration) and random models (model II, where the parents are a sample of genotypes from a reference population). Model I estimates apply only to the genotypes included and cannot be extended to some hypothetical reference population. Model II estimates are interpreted relative to some reference population from which the genotypes included are an unselected sample. The use of model I or II depends on sample size and this will vary among species (e.g., we could represent the tobacco species with 5–10 lines and the diallel mating design could be useful. Although limited sample sizes in some crops do not allow the estimation of heritability, genetic gain, genetic correlations with model I, we can get as much information as model II (GCA, SCA effects).

An orthogonal subdivision of the sums of squares for crosses is valid when the crosses mean square is significantly different from zero. An example of a diallel mating design (method IV or one set of crosses) including five parent and the partition of the cross sums of squares is shown in Tables 4.17 and 4.18, respectively. Obviously, a diallel of five parents does not apply to the model II analysis. However, assume only the 10 cross-combinations are evaluated in five replications. Table 4.17 shows that each parent is included in a cross with each of the others – four specific crosses, each including a common parent. The average performance of each parent in the four crosses is determined from the marginal means. Margins represent the average performance of parents in crosses or general combining ability (GCA) of parents. They are the deviations of the marginal means from the overall mean. On the other hand, cells represent the deviation of individual crosses from the average of the margins or specific combining ability (SCA) of parents. For model I, estimation of components of variance is not adequate but estimation of GCA and SCA effects is valid. This model is very useful for choice of parents (GCA) and/or hybrids (SCA). Hybrid seed companies are always searching for the best cell (SCA).

Hence total variation among crosses (variation among all cells or 9 df) can be partitioned into variation among margins (4 df for the five parents included) and variation among cells within margins or of individual cells about the margins, which have a common parent. There are 5 df for the variation among cells within margins because we have five independent observations among the cells in Table 4.17; i.e.,
Table 4.17 Example of the 10 cross-combinations possible from the diallel crossing of five parents

<table>
<thead>
<tr>
<th>Parents</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Margins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X_{12}</td>
<td>X_{13}</td>
<td>X_{14}</td>
<td>X_{15}</td>
<td>X_{11}</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>X_{23}</td>
<td>X_{24}</td>
<td>X_{25}</td>
<td>X_{22}</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>X_{34}</td>
<td>X_{35}</td>
<td>X_{33}</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>X_{45}</td>
<td>X_{44}</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X_{55}</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.18 Orthogonal partition of the crosses sums of squares in the analysis of variance of the diallel cross shown in Table 4.17

<table>
<thead>
<tr>
<th>SOV^b</th>
<th>df^a</th>
<th>Example</th>
<th>MS Model I</th>
<th>E(\text{MS})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r−1</td>
<td>9</td>
<td>\sigma^2 + rK^2</td>
<td>\sigma^2 + r\hat{\sigma}_g^2</td>
</tr>
<tr>
<td>Among m</td>
<td>\frac{n(n−1)}{2}−1</td>
<td>M_{21}</td>
<td>\sigma^2 + \frac{r(n−2)}{(n−1)}K^2_\delta</td>
<td>\sigma^2 + r\hat{\sigma}_g^2 + r(n−2)\hat{\sigma}_g^2</td>
</tr>
<tr>
<td>Among c/m</td>
<td>\frac{n(n−3)}{2}</td>
<td>M_{32}</td>
<td>\sigma^2 + \frac{2r}{n(n−3)}K^2_\delta</td>
<td>\sigma^2 + r\hat{\sigma}_g^2</td>
</tr>
<tr>
<td>Error</td>
<td>\frac{r(n(n−1)}{2}−1</td>
<td>M_{1}</td>
<td>\hat{\sigma}_g^2</td>
<td>\hat{\sigma}_g^2</td>
</tr>
<tr>
<td>Total</td>
<td>r\frac{n(n−1)}{2}−1</td>
<td>49</td>
<td>\hat{\sigma}_g^2</td>
<td>\hat{\sigma}_g^2</td>
</tr>
</tbody>
</table>

^a r and n refer to the number of replications and parents, respectively
^b m and c refer to margins and cells, respectively

If we know the marginal values, three cell values for parent 1, and two cell values for parent 2, the remaining cell values may be calculated. Thus the degrees of freedom for variation among cells within margins become 10 − 5 = 5. F-tests can be made to determine whether variation among margins and cells within margins are significantly different from zero.

Sprague and Tatum (1941) introduced the concepts of GCA and SCA to distinguish between the average performance of parents in crosses (GCA) and the deviation of individual crosses from the average of the margins (SCA). The concepts of GCA and SCA are extensively used in plant breeding and have particular significance to the diallel mating design. If we insert the GCA and SCA terms in Table 4.18, we have Table 4.19.

F-tests for both models can be made to test for GCA and SCA mean squares. For model I, \( M_{21} \) and \( M_{22} \) are tested with \( M_1 \); whereas for model II, \( M_{22} \) is tested with \( M_1 \) and \( M_{21} \) with \( M_{22} \). If \( M_{22} \), however, is not different from zero, \( M_1 \) also is the appropriate mean square to test \( M_{21} \).

In model I the parents are the population, and estimation of components of variance is not appropriate but estimation of the effects of each pair of parents for specific crosses (SCA) and for all crosses that include a common parent (GCA) is appropriate and valid. No apologies are needed from the experimenter for the estimation of effects rather than of components of variance; GCA and SCA effects
4.7 Diallel Methods

Table 4.19 Diallel analysis of variance for a fixed (model I) and random set (model II) of \( n \) parents to produce the \( n(n-1)/2 \) crosses

<table>
<thead>
<tr>
<th>Replications</th>
<th>df(^a)</th>
<th>MS</th>
<th>Model I</th>
<th>Model II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crosses</td>
<td>( [n(n-1)/2]-1 )</td>
<td>( M_2 )</td>
<td>( \hat{\sigma}^2 + rK^2_{GCA} )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_{GCA} )</td>
</tr>
<tr>
<td>GCA</td>
<td>( n-1 )</td>
<td>( M_{21} )</td>
<td>( \hat{\sigma}^2 + [r(n-2)/(n-1)]K^2_{GCA} )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_{GCA} + r(n-2)\hat{\sigma}^2_{GCA} )</td>
</tr>
<tr>
<td>SCA</td>
<td>( n(n-3)/2 )</td>
<td>( M_{22} )</td>
<td>( \hat{\sigma}^2 + (2r/[n(n-3)])K^2_{SCA} )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_{SCA} )</td>
</tr>
<tr>
<td>Error</td>
<td>( (r-1)[n(n-1)/2]-1 )</td>
<td>( M_1 )</td>
<td>( \hat{\sigma}^2 )</td>
<td>( \hat{\sigma}^2 )</td>
</tr>
<tr>
<td>Total</td>
<td>( n(n-1)/2-1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\( r \) and \( n \) refer to the number of replications and parents, respectively

The model for the analysis of variance is

\[ X_{ijk} = u + r_k + g_i + g_j + s_{ij} + p_{ijk} \]

where \( u \) is the mean, \( r_k \) is the replication effect, \( g_i \) and \( g_j \) are the GCA effects, \( s_{ij} \) is the SCA effect, and \( p_{ijk} \) is the experimental error for the \( X_{ijk} \) observation \((k = 1, 2, \ldots, r; i = j = 1, 2, \ldots, n)\). For model I, GCA and SCA effects are estimated, respectively, as

\[ \hat{\sigma}^2 = \frac{1}{\left[ n(n-1) \right]} \left( nX_i - 2X_.. \right) \]

\[ \hat{s}_{ij} = X_{ij} - \left[ 1/(n-2) \right] (X_i + X_j) + \left\{ 2/\left[ (n-1)(n-2) \right] \right\} X_.. \]

From Table 4.18, the experimental error \( \hat{\sigma}^2 \) is estimated from the mean square \( M_1 \). The variance of any specific \( X_{ij} \) cross is \( \hat{\sigma}^2/r \), the variance of the difference between any two crosses is \( 2\hat{\sigma}^2/r \), and the variance for the mean of the crosses where one parent is common is \( \hat{\sigma}^2/\left[ (n-1)r \right] \). The variances of the GCA and SCA effects are

\[ \hat{\sigma}^2(g_i) = \left\{ (n-1)/\left[ n(n-2) \right] \right\} \hat{\sigma}^2 \] and \( \hat{\sigma}^2(s_{ij}) = \left\{ (n-3)/(n-1) \right\} \hat{\sigma}^2 \)

The model I analysis, therefore, yields considerable information about the fixed set of parents included. This information that can be useful for the selection of parents that have good general combining ability in a series of crosses and good specific
combining ability for specific pairs of parents. This type of information is quite useful to maize breeders, particularly if the selected set of parents represents an elite group of inbred lines that are possible candidates as parent seed stock for the production of single-cross hybrids. Also, pre-breeding improved sources can be identified as new heterotic patterns for reciprocal recurrent selection programs.

For the model II analysis, estimation of the components of variance is of prime interest. To relate the variance components in Table 4.19 to types of gene action in the reference population, it is helpful to write the expected mean squares in terms of genetic relationships of relatives and translate from the covariances of relatives to the genetic components of variance (Table 4.20). Then it is necessary to determine what types of relatives are included in the diallel mating design. The variation among margins (GCA) is due to differences among parents what is equal to the covariance of progenies (crosses with a common parent).

The variance among margins (GCA) is due to differences among parents what is equal to the covariance of progenies (crosses with a common parent).

Therefore,

\[ \hat{\sigma}^2_{GCA} = \hat{\text{Cov}} \text{ HS} = \left(\frac{1}{4}\right)\hat{\sigma}^2_A \quad \text{if } F = 0 \]
\[ = \left(\frac{1}{2}\right)\hat{\sigma}^2_A \quad \text{if } F = 1 \]

The variance among cells (\(\hat{\sigma}^2_{SCA}\)) involves full-sib individuals. Therefore,

\[ \hat{\sigma}^2_{SCA} = \hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HS} = \left(\frac{1}{4}\right)\hat{\sigma}^2_D \quad \text{if } F = 0 \]
\[ = \hat{\sigma}^2_D \quad \text{if } F = 1 \]

### Table 4.20

<table>
<thead>
<tr>
<th>SOV</th>
<th>df(^a)</th>
<th>MS</th>
<th>EMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>(r-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crosses</td>
<td>([n(n-1)/2]-1)</td>
<td>(M_2)</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{GCA})</td>
</tr>
<tr>
<td>GCA</td>
<td>(n-1)</td>
<td>(M_{21})</td>
<td>(\hat{\sigma}^2 + r(\hat{\text{Cov}} \text{ FS} - 2\hat{\text{Cov}} \text{ HS}) + n(n-2)\hat{\text{Cov}} \text{ HS})</td>
</tr>
<tr>
<td>SCA</td>
<td>(n(n-3)/2)</td>
<td>(M_{22})</td>
<td>(\hat{\sigma}^2 + r(\hat{\text{Cov}} \text{ FS} - 2\hat{\text{Cov}} \text{ HS}))</td>
</tr>
<tr>
<td>Error</td>
<td>((r-1)([n(n-1)/2]-1))</td>
<td>(M_1)</td>
<td>(\hat{\sigma}^2_{b})</td>
</tr>
<tr>
<td>Total</td>
<td>(r[n(n-1)/2]-1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\(r\) and \(n\) refer to the number of replications and parents, respectively

\(^b\)If individual plant data are taken, \(\sigma^2\) is equal to \([(\hat{\sigma}^2_G - \hat{\text{Cov}} \text{ FS}) + \hat{\sigma}^2_{we}]/k + \hat{\sigma}^2_{p}\), where \(k\) is the number of plants.
Because we know that plot data, we would have another source of information relative to genetic variance. \( \hat{\sigma}^2 \), translations to genetic variances, i.e., shown in Table 4.20.

The relationship of relatives we can then write the model II analysis of variance as component mean of a particular parent. Next, let us consider sources of variation within the HS) because one parent is common for all the crosses included in the marginal mean of a particular parent. Then, let us consider sources of variation within each parent. Variability among i parents can be shown to be

\[
\hat{\sigma}^2_{y_i} = E_i(X_i - \bar{X})^2 = E_i(X_i^2) - \bar{X}^2
\]

If we have j individuals within families of the common i parent, covariance is

\[
\hat{\text{Cov}}(y_{ij}, y_{ij}') = E[(X_{ij} - \bar{X})(X_{ij'} - \bar{X})] = E(X_{ij}X_{ij'} - \bar{X}^2) = E(X_i^2) - \bar{X}^2
\]

because, \( E(X_{ij}) = E(X_{ij'}) = X_i \) and \( E(X_{ij}X_{ij'}) = X_i^2 \), assuming symmetrical distribution of sibs around family mean. Hence, \( \hat{\sigma}^2_{\text{GCA}} \) is equal to covariance of half-sibs (\( \hat{\text{Cov}} \text{ HS} \)) because one parent is common for all the crosses included in the marginal mean of a particular parent. Next, let us consider sources of variation within the component \( \hat{\sigma}^2 \), which is estimated from \( M_1 \) (Table 4.18). Experimental error (\( \hat{\sigma}^2 \)) includes a plot error variance (\( \hat{\sigma}^2_p \)) and variance among individuals within plots (\( \hat{\sigma}^2_w \)). Therefore, \( \hat{\sigma}^2 = \hat{\sigma}^2_w/k + \hat{\sigma}^2_p \), where \( k \) is the number of individuals measured within each plot. Variance within plots includes plant-to-plant environmental variance (\( \hat{\sigma}^2_w \)) and genetic variance (\( \hat{\sigma}^2_{\text{we}} \)) among individuals if \( F \neq 1 \); hence, \( \hat{\sigma}^2_w = \hat{\sigma}^2_{\text{we}} + \hat{\sigma}^2_{\text{we}} \). Individuals having the same parents are full-sibs, and consequently variations among individuals having the same parents are covariances of full-sibs (\( \hat{\text{Cov}} \text{ FS} \)). If individual plant data are taken, within-progeny genetic variance can be expressed as

\[
\hat{\sigma}^2_{\text{we}} = \hat{\sigma}^2_{\text{G}} - \hat{\text{Cov}} \text{ FS}, \quad \text{where} \quad \hat{\sigma}^2_{\text{G}} = \hat{\sigma}^2_{\text{A}} + \hat{\sigma}^2_{\text{D}} + \hat{\sigma}^2_{\text{AA}} + \cdots, \quad \text{total genetic variance.}
\]

Then, \( \hat{\sigma}^2_{\text{we}} = \hat{\sigma}^2_{\text{G}} - \hat{\text{Cov}} \text{ FS} = (\gamma_2)\hat{\sigma}^2_{\text{A}} + (\gamma_4)\hat{\sigma}^2_{\text{D}} + (\gamma_4)\hat{\sigma}^2_{\text{AA}} + \cdots \) for \( F = 0 \).

Since, \( \hat{\text{Cov}} \text{ HS} \) is \( (\gamma_4)\hat{\sigma}^2_{\text{A}} + (\gamma_4)\hat{\sigma}^2_{\text{AA}} + \cdots \) and \( \hat{\sigma}^2_{\text{we}} \) is \( (\gamma_2)\hat{\sigma}^2_{\text{A}} + (\gamma_4)\hat{\sigma}^2_{\text{D}} + (\gamma_4)\hat{\sigma}^2_{\text{AA}} + \cdots \) of total genetic variance (\( F = 0 \)) is accounted for. Remaining genetic variation is among crosses within each parent. The crosses within each parent are half-sibs but consist of full-sib individuals. We can determine variation among crosses for each parent, but again this does not tell us anything about genetic relationships. The crosses consist of full-sib individuals, and therefore the variance among cross means is equal to \( \hat{\text{Cov}} \text{ FS} - \hat{\text{Cov}} \text{ HS} \), where we subtract 2 \( \hat{\text{Cov}} \text{ HS} \), because each of the parents is included in each cross. From the relationship of relatives we can then write the model II analysis of variance as shown in Table 4.20.

We have two types of relatives in the diallel mating design for use in making translations to genetic variances, i.e., \( \hat{\text{Cov}} \text{ HS} \) and \( \hat{\text{Cov}} \text{ FS} \). If we obtained within-plot data, we would have another source of information relative to genetic variance. Because we know that

\[
\hat{\text{Cov}} \text{ HS} = (\gamma_4)\hat{\sigma}^2_{\text{A}} + (\gamma_4)\hat{\sigma}^2_{\text{AA}} + \cdots \quad \text{and} \quad \hat{\text{Cov}} \text{ FS} = (\gamma_2)\hat{\sigma}^2_{\text{A}} + (\gamma_4)\hat{\sigma}^2_{\text{D}} + (\gamma_4)\hat{\sigma}^2_{\text{AA}} + \cdots
\]

for \( F = 0 \), we can use these to obtain \( \hat{\sigma}^2_{\text{A}} \) and \( \hat{\sigma}^2_{\text{D}} \).

Assuming no epistasis,

\[
\hat{\sigma}^2_{\text{A}} = 4 \hat{\text{Cov}} \text{ HS} = 4(M_{21} - M_{22})/[r(n - 2)]
\]
and

\[ \hat{\sigma}_D^2 = 4\left[ \hat{\text{Cov}} \, FS - 2 \hat{\text{Cov}} \, HS \right] = 4 (M_{22} - M_1)/r \]

If we collect individual plant data and have \( \hat{\sigma}_{\text{we}}^2 \), we can obtain \( \hat{\sigma}_{\text{wg}}^2 \), which is \( (\gamma_2) \hat{\sigma}_A^2 + (\gamma_4) \hat{\sigma}_D^2 \).

A rough check on the assumption of no epistasis then can be tested as

\[ \hat{\sigma}_{\text{wg}}^2 - 3 \hat{\text{Cov}} \, FS + 4 \hat{\text{Cov}} \, HS \]

If epistasis is either absent or relatively small, this relation will be zero or some relatively small value. Otherwise, the presence of significant epistasis will cause this relation to be greater than zero because the difference of the relation theoretically includes the following relative proportions of the various types of epistasis:

\[ (\gamma_4) \sigma_{AA}^2 + (\gamma_2) \sigma_{AD}^2 + (\gamma_4) \sigma_{DD}^2 + \cdots \]

This test, however, has two major disadvantages: (1) the great variability of individual plant data and (2) the large errors that are associated with such estimates obtained from complex linear functions.

Quite often inbred parents \( (F = 1) \) are used, such as inbred lines of maize and cultivars of self-pollinated crop species. The model I analysis would be appropriate for selected, elite inbred lines, and no complications are encountered estimating the GCA and SCA effects. This information is often very useful to the maize breeder.

The inclusion of inbred lines as parents for the model II analysis, however, does change the coefficients of the components of genetic variance of the covariances of relatives. Adding epistasis estimates we find that for \( F = 1 \) we have the following:

\[ \hat{\text{Cov}} \, HS = (\gamma_2) \hat{\sigma}_A^2 + (\gamma_4) \hat{\sigma}_{AA}^2 + \cdots \]

and

\[ \hat{\text{Cov}} \, FS = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{AA}^2 + \cdots \]

Hence, \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \), assuming no epistasis, are as follows:

\[ \hat{\sigma}_A^2 = 2 \hat{\text{Cov}} \, HS \] and \[ \hat{\sigma}_D^2 = \hat{\text{Cov}} \, FS - 2 \hat{\text{Cov}} \, HS \]

Individual plant data collected on crosses of inbred lines would include only \( \hat{\sigma}_{\text{we}}^2 \), which is within-plot environmental variance. The only restriction in estimating components of genetic variance from the model II analysis that uses inbred lines as parents is that inbred lines are an unselected sample from the reference population to which \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) are to be applied.
If we assume that the parents are an unselected sample from some hypothetical population (model II, Table 4.19), estimates of heritability on cross means could be calculated as follows:

\[ \hat{h}^2 = \frac{\hat{\sigma}_c^2}{\hat{\sigma}^2 / r + \hat{\sigma}_c^2} \]

Within environments

which is the same as with bi-parental progenies (Tables 4.1 and 4.2).

We have so far considered only one of the possible diallel mating systems (e.g., one set of crosses). Other possible designs include the combination of crosses and parents and the permutations of crosses with or without parents (Griffing, 1956). Combinations of crosses without parents is probably the most commonly used method in maize because parents are usually inbred lines and the vigor of parents \((F = 1)\) and crosses among parents \((F = 0)\) frequently cause complications in field designs used to evaluate parents and crosses. Border rows are necessary if mature plant traits are measured and differences in seedling vigor cause problems in measuring traits in the juvenile stage of development. However, the method that includes parents and crosses are common for non-inbred parents too as explained later on (e.g., improved populations and their hybrids). Cockerham (1963) presented an analysis for partitioning the cross sums of squares for GCA, SCA, maternal, and reciprocal sources of variation when all possible permutations among parents are included especially used for traits influenced by maternal and reciprocal effects (e.g., Jumbo and Carena, 2008).

If parents are included for evaluation, the source of variation of parents vs. crosses provides another test (in addition to SCA) for the importance of non-additive effects. If we assume the parents and all combinations of crosses (no reciprocals) are included for analysis, we can orthogonally partition variation among entries as shown in Table 4.21. For the model I analysis, tests can be made for variation among parents and parents vs. crosses, and among crosses. Partitioning of the crosses sums of squares would be the same as shown in Table 4.20. The analysis given in Table 4.20 is the same as Griffing’s experimental method 2 except that SCA has \(n(n - 3)/2\) df rather than the \(n(n - 1)/2\) df given by Griffing. The difference of \(n\) df is for the separation of among parents and parents vs. crosses from the SCA source of variation. The parents vs. crosses comparison is a test for heterosis, which also is due to non-additive genetic effects.

Diallel analyses that include the parents are frequently used for open-pollinated, synthetic, and composite varieties either improved or not. For these instances we are interested in the variety performance itself as well as the variety crosses. Border rows on the plots may not be needed because the differences in vigor are usually small. Variety and variety cross evaluations are important in maize breeding for (1) determining the relative potential of varieties as breeding populations and (2) evaluating the response of varieties to different recurrent selection schemes. Because the varieties usually are a selected sample of the most promising available, the model I analysis is appropriate to determine the GCA effects of the varieties and SCA effects of the variety crosses.
Table 4.21 Diallel analysis that includes n parents and their \( n(n-1)/2 \) crosses

<table>
<thead>
<tr>
<th>SOV</th>
<th>df(^a)</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>( r-1 )</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Entries</td>
<td>([n(n+1)/2]-1)</td>
<td>54</td>
<td>( S_2 )</td>
</tr>
<tr>
<td>Parents (P)</td>
<td>( n-1 )</td>
<td>9</td>
<td>( S_{21} )</td>
</tr>
<tr>
<td>P vs. C</td>
<td>1</td>
<td>1</td>
<td>( S_{22} )</td>
</tr>
<tr>
<td>Crosses (C)</td>
<td>([n(n-1)/2]-1)</td>
<td>44</td>
<td>( S_{23} )</td>
</tr>
<tr>
<td>GCA</td>
<td>( n-1 )</td>
<td>9</td>
<td>( S_{231} )</td>
</tr>
<tr>
<td>SCA</td>
<td>( n(n-3)/2 )</td>
<td>35</td>
<td>( S_{232} )</td>
</tr>
<tr>
<td>Error</td>
<td>((r-1)[n(n+1)/2]-1)</td>
<td>108</td>
<td>( S_1 )</td>
</tr>
</tbody>
</table>

\(^a\)\( r \) and \( n \) refer to the number of replications and parents, respectively

4.7.1 The Gardner–Eberhart Analysis II

Gardner and Eberhart’s (1996) analysis II is especially useful for evaluating varieties and their crosses. The Gardner–Eberhart analysis II model includes the \( n \) parent varieties and their \( n(n-1)/2 \) variety crosses, but the partitioning of the entry sums of squares differs from the analysis given in Table 4.21. The following models are used to determine the sums of squares for the analysis shown in Table 4.22:

1. \( X_{ij}' = u + (\frac{1}{2})(v_j + v'_j) = (B'G)_1 \)
2. \( X_{ij}' = u + (\frac{1}{2})(v_j + v'_j) + vH = (B'G)_2 \)
3. \( X_{ij}' = u + (\frac{1}{2})(v_j + v'_j) + vH + v(h_j + h'_j) = (B'G)_3 \)
4. \( X_{ij}' = u + (\frac{1}{2})(v_j + v'_j) + vH + v(h_j + h'_j) + v s_{jj}' = (B'G)_4 \)

Table 4.22 Analysis of variance of \( n \) parents and their \( n(n-1)/2 \) variety crosses for variety and heterosis effects (analysis II of Gardner–Eberhart model)

<table>
<thead>
<tr>
<th>SOV</th>
<th>df(^a)</th>
<th>SS</th>
<th>Gardner–Eberhart Diallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>( r-1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entries</td>
<td>([n(n+1)/2]-1)</td>
<td>( S_1 )</td>
<td>( M_1 )</td>
</tr>
<tr>
<td>Varieties (( v_i ))</td>
<td>( n-1 )</td>
<td>( S_{21} )</td>
<td>( = (B'G)_1 - CF )</td>
</tr>
<tr>
<td>Heterosis (( h_{ij} ))</td>
<td>( n(n-1)/2 )</td>
<td>( S_{22} )</td>
<td>( = (B'G)_4 - (B'G)_1 )</td>
</tr>
<tr>
<td>Average (( \bar{H} ))</td>
<td>1</td>
<td>( S_{231} )</td>
<td>( = (B'G)_2 - (B'G)_1 )</td>
</tr>
<tr>
<td>Variety (( h_i ))</td>
<td>( n-1 )</td>
<td>( S_{232} )</td>
<td>( = (B'G)_3 - (B'G)_2 )</td>
</tr>
<tr>
<td>Specific (( s_{ij} ))</td>
<td>( n(n-3)/2 )</td>
<td>( S_{233} )</td>
<td>( = (B'G)_4 - (B'G)_3 )</td>
</tr>
<tr>
<td>Error</td>
<td>((r-1)[n(n+1)/2]-1)</td>
<td>( S_1 )</td>
<td>( = M_1 )</td>
</tr>
</tbody>
</table>

\(^a\)\( r \) and \( n \) refer to the number of replications and parents, respectively
In each of the models, $u$, $v_j$, $h$, $h_j$, and $sjj$ indicate the mean and variety and heterosis effects. The coefficient $v$ in these models is zero when $j = j'$ and one when $j \neq j'$. Because the phenomenon of heterosis is important, the analysis maximizes the information on variety performance and the expression of heterosis of their crosses. Estimates of the variety and heterosis effects can be determined for each of the constants in the models. Four of the mean squares of the diallel analysis (Table 4.21) and analysis II of the Gardner–Eberhart model (Table 4.22) are equivalent; i.e., entry and error mean squares are the same, average heterosis mean square is equal to the parents vs. crosses mean square, and specific heterosis mean square is equal to the SCA mean square. The variety heterosis mean square in Table 4.22 is not equivalent to the parent mean square in Table 4.21 because the variety heterosis mean square includes information of the performance of varieties themselves and in variety crosses. Also, the GCA mean square (Table 4.21) is not equivalent to the variety heterosis (Table 4.22) because $gi = (1/2)v_j + h_j$ so that $M_{21} + M_{231} = M_{21}^1 + M_{222}^1$. The analysis in Table 4.22 is a non-orthogonal partition of the entry sums of squares, but the sums of squares can be obtained by sequentially fitting the four models.

An analysis of the diallel mating design repeated over environments (e.g., partitioned in years and locations) was given by Matzinger et al. (1959). Nothing changes in regard to the assumptions regarding parents relative to model I and model II analyses except that estimates of interaction of effects (model I) and variances (model II) with environments can be made. An example of a diallel mating design including crosses and parents repeated over environments is shown in Table 4.23.

<table>
<thead>
<tr>
<th>Table 4.23 Diallel analysis of $n(n+1)/2$ parents and their crosses repeated over environments</th>
</tr>
</thead>
<tbody>
<tr>
<td>df$^a$</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>General</td>
</tr>
<tr>
<td>Environments (E)</td>
</tr>
<tr>
<td>Replications/E</td>
</tr>
<tr>
<td>Entries</td>
</tr>
<tr>
<td>Parents (P)</td>
</tr>
<tr>
<td>P vs. C</td>
</tr>
<tr>
<td>Crosses (C)</td>
</tr>
<tr>
<td>GCA</td>
</tr>
<tr>
<td>SCA</td>
</tr>
<tr>
<td>E x entries</td>
</tr>
<tr>
<td>E x P</td>
</tr>
<tr>
<td>E x P vs. C</td>
</tr>
<tr>
<td>E x C</td>
</tr>
<tr>
<td>E x GCA</td>
</tr>
<tr>
<td>E x SCA</td>
</tr>
<tr>
<td>Pooled error</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

$^a$e, r, and n refer to the number of environments, replications, and parents, respectively.
If we assume environmental effects are random, F-tests of main effects use the same mean squares for models I and II. If environmental effects are fixed, as are parents in model I, all main effects and their interactions with environments are tested with the pooled error for model I.

For the model II analysis, estimates of components of genetic variance permit estimation of heritability ($h^2$). From Table 4.20 the source of variation due to GCA is the covariance of half-sibs, or $(\hat{\sigma}_A^2)$, and the variation due to SCA is the covariance of full-sibs minus two times the covariance of half-sibs, or $(\hat{\sigma}_D^2)$, assuming no inbreeding of the parents and no epistasis. Using the components of variance shown in Table 4.19, heritability based on the means of r plots (entry-mean basis) can be calculated as

$$\hat{h}^2 = \frac{4\hat{\sigma}_{GCA}^2}{(\hat{\sigma}_2^2 / r + 4\hat{\sigma}_{SCA}^2 + 4\hat{\sigma}_{GCA}^2)}$$

which is commonly referred to as heritability in the narrow sense. Standard errors of the heritability estimate can be calculated from the variance of a ratio of two sets of variance components, as illustrated for bi-parental progenies. Dickerson (1969, pp. 36–79) presented a simplified general formula for calculating conservative approximations of the standard errors of heritabilities, based on the methods of Graybill et al. (1956) and Graybill and Robertson (1957). The simplified formula is

$$\hat{\sigma}(X/Y) = (C/Y)[V(\hat{X})]^{1/2}$$

which neglects the terms involving $V(\hat{Y})$ and $\text{Cov}(\hat{X}, \hat{Y})$. In the formula for the estimate of heritability, $\hat{\sigma}_{GCA}^2$ is equivalent to $\hat{X}$ and the total, or phenotypic, variance is equivalent to $\hat{Y}$. Hence the standard error (SE) of our estimate of heritability becomes

$$\text{SE}(\hat{h}^2) = 4\text{SE}(\hat{\sigma}_{GCA}^2) / (\hat{\sigma}_2^2 / r + 4\hat{\sigma}_{SCA}^2 + 4\hat{\sigma}_{GCA}^2)$$

The variance of the estimate of $\hat{\sigma}_{GCA}^2$ is obtained in the usual manner from Table 4.19 as

$$\frac{2}{[r(n - 2)]^2} \left[ \frac{M_{21}^2}{n + 1} + \frac{M_{22}^2}{n(n - 3)/2 + 2} \right]$$

Estimates of components of variance for one environment include a genotype–environment bias, so that $(\hat{\sigma}_A^2 + \hat{\sigma}_{AE}^2)$ is estimated by $\hat{\sigma}_A^2$. If diallel experiments are repeated over environments, an estimate of $\hat{\sigma}_A^2$ unbiased by environmental interactions can be obtained, and the phenotypic variance would include additional terms due to the interaction of $\hat{\sigma}_{GCA}^2$ and $\hat{\sigma}_{SCA}^2$ with environments. Estimates of heritability from the diallel mating design are only as good as estimates of components of variance. Because of the number of parents that can be included in the diallel, standard errors of the components of variance may be quite large, particularly for the GCA
source of variation. Standard errors on estimates of heritability, consequently, may be large as well.

Sokol and Baker (1977) and Baker (1978) have reviewed the critical issues involved in the use of the diallel mating design. The critical issues concern the choice of model for analysis of data, i.e., model I (fixed genotypic effects) or model II (random genotypic effects). Baker (1978) emphasizes that two assumptions are critical for interpreting the results of the diallel analyses: (1) independent distribution of genes in the parents included and (2) no epistasis. Neither assumption seems valid for the small number of parents usually included in a diallel set of crosses. Independent distribution of genes at $n$ loci cannot occur unless a minimum of $2^n$ parents are included for the diallel set of crosses. The assumption of no epistasis is commonly made, and estimation of the relative importance of epistasis usually has not been fruitful (Chapter 5) even at the molecular level. But epistatic effects unpredictably affect the GCA and SCA mean squares, variances, and effects (Baker, 1978). He concluded that most diallel experiments are restricted to estimation of GCA and SCA mean squares and effects.

Use of diallel mating designs has been common in self-fertilizing species. To overcome the problem of limited hybrid ($F_1$) seed often experienced from crosses of pure-line cultivars, $F_1$ generation seed may be advanced one generation to produce $F_2$ generation seed so as to have adequate seed for testing. The analyses of diallel mating designs discussed so far included non-inbred progeny produced from either non-inbred or inbred (or partially inbred) parents, i.e., the progeny evaluated was non-inbred but parents may have some level of inbreeding or none. Use of $F_2$ generation (or any other advanced generation obtained by inbreeding) seed in the evaluation trials of crosses among parents means the progenies evaluated have some level of inbreeding. Stuber (1970) examined the situations that included evaluation of inbred progenies where they are generated by bulk selfing of each $F_1$ cross. The main effect of evaluating inbred progenies is the change in coefficient of the dominance variance. Although the evaluation of inbred progenies of a cross is applicable for both cross- and self-fertilizing species, its greatest usefulness would be for self-fertilizers. Stuber (1970) also discussed how the evaluation of progenies with different levels of inbreeding that are developed from parents with different levels of inbreeding may be used in the estimation of genetic components of variance.

4.8 Partial Diallel

The partial diallel design was developed by Kempthorne and Curnow (1961). This design is a modification of the diallel with the purpose to increase the number of parents that can be included in a diallel mating design. The mechanical procedures for developing crosses and the principles of analyzing data are similar to those of the complete diallel. The major difference between partial and complete diallels is the number of crosses made among parents. For the diallel, $n(n - 1)/2$ combinations of crosses are made among $n$ parents, whereas for the partial diallel less crosses are
made. So, one important modification is the number of crosses made among parents. Therefore, an important advantage over the diallel is the estimation of genetic variances with a fixed number of resources (greater number of parents).

The partial diallel has a total of $ns/2$ crosses, where $n$ is the number of parents, $s$ is a whole number greater than or equal to 2, and $k$ is a whole number ($k = (n+1-s)/2$). If we have $n$ parents, the following crosses can be sampled:

\[
\begin{align*}
1 \times (k + 1) & \quad 2 \times (k + 2) & \cdots & \quad n \times (k + n) \\
1 \times (k + 2) & \quad 2 \times (k + 3) & \cdots & \quad n \times (k + n + 1) \\
\vdots & \quad \vdots & \cdots & \quad \vdots \\
1 \times (k + s) & \quad 2 \times (k + 1 + s) & \cdots & \quad n \times (k + n - 1 + s)
\end{align*}
\]

For $k$ to be a whole number, we do not want both $n$ and $s$ to be odd or both even. Each parent occurs in $s$ crosses, and the number of crosses sampled is $ns/2$. If $s = n - 1$, it corresponds to the complete diallel of $n(n - 1)/2$ cross-combinations.

As an example, assume we have the resources to make and grow 120 crosses. If $ns/2 = x$, we can determine the number of parents that can be included for number of crosses $s$, i.e., $n = 2x/s$. Table 4.24 summarizes relations among mating designs for the number of parents that can be included if 120 crosses are made and evaluated.

**Table 4.24** Number of parents that can be included to produce 120 crosses for the diallel, partial diallel, and design II mating designs

<table>
<thead>
<tr>
<th>Partial diallel</th>
<th>Design II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Parents ($n$) Crosses</td>
<td>$s = 3$ $s = 4^a$ $s = 5$</td>
</tr>
<tr>
<td>Parents ($n$)</td>
<td>16 120</td>
</tr>
</tbody>
</table>

$^a$Not valid because for $k$ to be a whole number, $n$ and $s$ cannot both be even

We can include three (when $s = 5$) to five (when $s = 3$) times more parents in the partial diallel relative to the diallel when we consider the number of crosses is fixed at 120. Design II is similar to $s = 4$ (a non-valid $s$ value) when eight sets of eight parents are included.

The particular number of crosses made and tested among the $n$ parents is determined from the calculated $k$ value. If we assume $n = 80$ and $s = 3$, we will have $(80 \times 3)/2 = 120$ crosses, and $k = (80 + 1 - 3)/2 = 39$. Hence, the cross-combinations among the 80 parents will be

\[
\begin{align*}
1 \times 40 & \quad 2 \times 41 & \cdots & \quad 39 \times 78 & \quad 40 \times 79 & \quad 41 \times 80 & \quad 42 \times 1^a & \cdots & \quad 80 \times 39^a \\
1 \times 41 & \quad 2 \times 42 & \quad 39 \times 79 & \quad 40 \times 80 & \quad 41 \times 1^a & \quad 42 \times 2^a & \quad 80 \times 40^a \\
1 \times 42 & \quad 2 \times 43 & \quad 39 \times 80 & \quad 40 \times 1 & \quad 41 \times 2^a & \quad 42 \times 3^a & \quad 80 \times 41^a
\end{align*}
\]

$^a$A reciprocal – one or the other not be grown
Each parent, therefore, is equally represented in the crosses. The advantage of the partial diallel over other mating designs can be clearly seen when we use large samples to estimate genetic variances (Table 4.25).

The ANOVA is quite similar to the diallel mating design. The experimental design model, the genetic model, the analysis of variance, and the covariances of relatives of the partial diallel are the same as for the diallel. However, the degrees of freedom and coefficients of expected mean squares are different because of the sampling of crosses among parents (Table 4.26). Therefore, in addition to including a greater number of parents, the partial diallel also has the advantage of having a more even distribution of degrees of freedom for GCA and SCA (Table 4.27), because more parents are included for a given number of crosses. As a consequence, components of variance for GCA and SCA are obtained with similar precision while in the diallel the degrees of freedom are smaller for the GCA mean square in relation to the SCA mean square. This may be a serious disadvantage in the estimation of \( \hat{\sigma}^2_g \) unless \( \hat{\sigma}^2_s \) is small compared with \( \hat{\sigma}^2 + r \hat{\sigma}^2_s \). For \( s = 3 \) nearly twice as many degrees of freedom are included for GCA compared with SCA. When \( s = 5 \), the ratio of GCA to SCA df is 0.65. With a more even distribution of degrees of freedom, components of variance will be estimated with approximately the same precision.

### Table 4.25  Number of crosses needed to represent 100-parent maize populations for the diallel, design II, and partial diallel mating designs

<table>
<thead>
<tr>
<th></th>
<th>Diallel</th>
<th>Design II</th>
<th>Partial diallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lbrack n(n-1)/2 \rbrack )</td>
<td>50 × 50</td>
<td>ns/2</td>
<td></td>
</tr>
<tr>
<td>4,950</td>
<td>2,500</td>
<td>150 (if ( s=3 ))</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.26  Analysis of variance of the partial diallel evaluated in one environment, model II

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>( r-1^a )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crosses</td>
<td>( (ns/2)-1 )</td>
<td>( M_1 )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_s )</td>
</tr>
<tr>
<td>GCA</td>
<td>( n-1 )</td>
<td>( M_{31} )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_s + [rs(n-2)/(n-1)]\hat{\sigma}^2_g )</td>
</tr>
<tr>
<td>SCA</td>
<td>( n(s/2-1) )</td>
<td>( M_{32} )</td>
<td>( \hat{\sigma}^2 + r\hat{\sigma}^2_s )</td>
</tr>
<tr>
<td>Error</td>
<td>( (r-1)[(ns/2)-1] )</td>
<td>( M_2 )</td>
<td>( \hat{\sigma}^2_s )</td>
</tr>
<tr>
<td>Total</td>
<td>( (rns/2)-1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>( r(ns/2)(k-1) )</td>
<td>( M_1 )</td>
<td></td>
</tr>
</tbody>
</table>

*\( a \), \( r \), \( n \), \( s \), and \( k \) refer to the number of replications, parents, crosses per parent, and plants within a plot, respectively.
*If individual plant data are taken, \( \hat{\sigma}^2 \) will be equal to \( [(\hat{\sigma}^2_g - \text{Cov} \text{ FS}) + \hat{\sigma}^2_{\text{fw}}]/k + \hat{\sigma}^2_p \), where \( k \) is the number of plants measured per plot and \( \hat{\sigma}^2_p \) is the experimental plot error.
Table 4.27  Distribution of the degrees of freedom for the partial diallel and diallel for 120 crosses grown in two replications in one environment

<table>
<thead>
<tr>
<th>Degrees of freedom</th>
<th>Partial diallel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>1</td>
</tr>
<tr>
<td>Crosses</td>
<td>119</td>
</tr>
<tr>
<td>GCA</td>
<td>15</td>
</tr>
<tr>
<td>SCA</td>
<td>104</td>
</tr>
<tr>
<td>Error</td>
<td>119</td>
</tr>
<tr>
<td>Total</td>
<td>239</td>
</tr>
</tbody>
</table>

\( s = 4 \) is not valid because for \( k \) to be a whole number, \( n \) and \( s \) will not be even.

Kempthorne and Curnow (1961) also determined the relative efficiency of the partial diallel to the complete diallel by comparing the relative yielding capacities of the crosses through methods of estimation. Two methods of estimation were designated A and B: In A the unsampled crosses are estimated by \( \hat{u} + \hat{g}_i + \hat{g}_j \) but sampled crosses are estimated by cross means \( y_{ij} \); in B both sampled and unsampled crosses are estimated by \( \hat{u} + \hat{g}_i + \hat{g}_j \).

Method B is preferable to method A only if \( r \hat{\sigma}^2_s / \hat{\sigma}^2 \) is small. The methods are identical when \( s = 2 \), but \( s \) has to be greater than 2 for \( \hat{\sigma}^2 \) to be estimable. If \( s > 2 \), A is preferred to B only if \( \hat{\sigma}^2_s / (r \hat{\sigma}^2_s) < 1 \). If \( r > 1 \), a decision between A and B is determined by the estimate of \( \hat{\sigma}^2_s / (r \hat{\sigma}^2_s) \) from the analysis of variance.

A third comparison C was made by estimating performances of the crosses by crossing the parents with each of \( t \) common testers. When \( n = 5 \), the partial diallel with \( s = 2 \) is preferred to use of the common tester only if \( \hat{\sigma}^2_s / (r \hat{\sigma}^2_s) < \frac{1}{3} \). When \( n \geq 7 \), the use of one common tester is preferred to the partial diallel with \( s = 2 \) for all values of \( \hat{\sigma}^2_s / (r \hat{\sigma}^2_s) \).

Since its introduction the partial diallel has had limited testing (Jensen, 1959). For a small number of selected parents the partial diallel will not provide any more information than can be obtained from the complete diallel. For a larger number of parents, design II is simpler to use than the partial diallel. The partial diallel provides another alternative, but the design does not seem to have potential for extensive use.

The heritability of the cross means would be

\[
\hat{h}^2 = \frac{\hat{\sigma}_c^2}{\hat{\sigma}^2 / r + \hat{\sigma}_c^2} \quad \text{(from Table 4.26)}
\]

If the crosses were evaluated in different environments \( (e) \), the heritability of cross means would be

\[
\hat{h}^2 = \frac{\hat{\sigma}_c^2}{\hat{\sigma}^2 / re + \hat{\sigma}_{ce}^2 / e + \hat{\sigma}_c^2} \quad \text{(from Table 4.26)}
\]
The estimates of heritability can be used to predict genetic response to selection (see Chapter 6). Partial diallels are also used to test single crosses among selected inbred lines. Data from partial diallels are used in best linear unbiased prediction (BLUP) and best linear unbiased estimation (BLUE) analyses to predict the untested single crosses, which are tested in future trials (see Bernardo, 2002, pp. 211–238).

The partial diallel mating design would be a more appropriate mating design, in most instances, than the diallel for estimation of genetic components of variance with similar accuracy (e.g., GCA and SCA) because a greater number of parents can be included, for the resources available, to have a better sample that is representative of the genetic variation of the population.

### 4.9 Triple Testcross

Kearsey and Jinks (1968) developed the triple testcross (TTC) design and analysis as an extension of Comstock and Robinson’s (1952) design III. The TTC can be used to detect epistatic effects for quantitative traits and provides estimates of additive and dominance genetic variances in the absence of epistasis.

![Diagram of triple testcross](image)

Similar to the design III mating design, individual $S_0$ plants in an $F_2$ population are used as males to cross to each parent ($P_1$ and $P_2$) and the $F_1$ ($P_1 \times P_2$) to produce three sets of testcrosses. For the $i$th male of the $F_2$ population, Kearsey and Jinks (1968) designated $L_{1i}$, $L_{2i}$, and $L_{3i}$ as the respective testcrosses of the $i$th male crossed to the $P_1$, $P_2$, and $F_1$ testers. They proposed the expression $L_{1i} + L_{2i} - 2L_{3i} = D$, the epistatic deviation. The epistatic deviation $D$ should equal zero in the absence of epistasis, and $D$ will significantly differ from zero if epistatic effects are present. When computing $D$, the additive and dominance terms cancel and the epistatic terms remain, which is true for any number of loci. Irrespective of the genetic constitution of the population (i.e., gene frequencies and linkage disequilibrium), the proposed method will test for net epistatic effects summed across all loci.
at which $P_1$ and $P_2$ differ. Perkins and Jinks (1970) showed that the TTC analysis provides two $F$-tests for the presence of epistatic effects. Genetic variation among the three testers was partitioned into two orthogonal contrasts testing for epistatic effects:

Contrast 1 = $L_1 + L_2 - 2L_3$. Test for the presence of additive by additive epistatic effects

Contrast 2 = $L_{1i} + L_{2i} - 2L_{3i}$. Test for additive by dominance + dominance by dominance

The $L_1 + L_2 - 2L_3$ contrast was designated as epistasis in the TTC analysis. The tester by male source of variation was partitioned into two sources of variation, one of which was the variation in $L_{1i} + L_{2i} - 2L_{3i}$ among males, which was designated as epistasis by male and tests for additive by dominance and dominance by dominance epistatic effects. Both sources of epistatic variation also could be tested for their interactions with environments if testcrosses are grown in multiple environments. A $t$-test was used to determine if deviation means ($\bar{D}$) were different from zero:

$$t = \frac{(\bar{D} - \mu_0)}{[V(\bar{D})/n]^{1/2}}$$

where $\mu_0 = 0$, $n$ is the number of observations in the mean, $V(\bar{D})$ is the variance ($V$) of $\bar{D}$, which is $V(L_{1i}) + V(L_{2i}) + 4V(L_{3i})$.

Wolf and Hallauer (1997) reported significant estimates of epistatic effects in the (B73 × Mo17) cross with the TTC design. The design III (DIII) and TCC mating designs were developed primarily to test for levels of dominance (DIII) and presence of epistasis (TTC). Although restricted to $F_2$ populations, estimates of components of genetic variance also can be calculated, provided adequate number of males (e.g., $n = 100$) are sampled from the $F_2$ populations. Wolf et al. (2000) also self-pollinated the $F_2$ males used to make the DIII crosses and evaluated the $S_1$ progenies and the two sets of testcrosses. From the DIII testcrosses, $S_1$ progenies, and the covariance analyses between DIII testcrosses and $S_1$ progenies, 10 mean squares and mean products were translated to genetic components of variance and error variances. The mean squares and mean products were used to estimate components of genetic variance through digenic epistatic components of variance. But Wolf et al. (2000) found that the estimates of epistatic variances generally were not significantly different from zero and were relatively less important than the estimates of the additive genetic and dominance components of variance for the (B73 × Mo17) $F_2$ population.

### 4.10 Triallel and Quadrangle

Cockerham (1961) determined the covariances, in terms of genetic components of variance, between all possible pairs of hybrid relatives among single-cross, three-way, and double-cross hybrids produced from a group of parents that originated
from the same population. Analysis of single-cross hybrids was shown in the various diallel analyses in which only additive and dominance genetic variances were estimable, with the assumption of no epistatic effects. Rawlings and Cockerham (1962a, b) presented the triallel (three-way crosses) and quadrallel (double-cross hybrids) analyses to clarify gene action involved in hybrids and to provide estimates of genetic components of variance and tests of genetic hypotheses. As in the model II analyses of the diallel it is important to emphasize that parents used to produce hybrids are randomly chosen from a reference population, which is necessary for estimation of genetic variance components and making tests of genetic hypotheses. Maize breeders usually classify their parent lines for use in hybrids and produce hybrids relative to unrelatedness of parent lines. This is valid, but the researcher should not attempt to estimate components of genetic variance and interpret estimates relative to some reference population. Hence for model II analyses of diallel, triallel, and quadrallel mating designs, the parents of the hybrids are not related because they originated from randomly chosen S0 plants in the reference population. The parents per se can be inbred lines, but they are not related in the sense of having common parents in their ancestry. It is assumed that all lines have the same level of inbreeding.

Because results of triallel and quadrallel analyses are to be interpreted relative to a particular reference population, it is necessary to sample the reference population adequately. If we have \( n \) lines,

\[
n(n - 1)(n - 2)/6, \text{ three-way cross-combinations are possible}
\]

The possible arrangements or permutations of the three crosses are

\[
3[n(n - 1)(n - 2)/6], \text{ assuming no reciprocal crosses}
\]

Similarly, if we have \( n \) parents, there are

\[
n(n - 1)(n - 2)(n - 3)/24 \text{ possible double crosses; and considering the three possible arrangements of four parents, we will have}
\]

\[
3[n(n - 1)(n - 2)(n - 3)/24] \text{ possible double crosses}
\]

Table 4.28 illustrates relative proportions of the three types of hybrids for \( n \) parents. Even for \( n = 10 \) the number of three-way and double crosses is large for testing. Fifty parents from a population is not unreasonable, but the number of crosses to test is unmanageable. The reduction in the number of crosses for testing (but including a reasonable number of parents to sample reference population genotypes) can be obtained by partitioning the parents into sets and pooling sums of squares and degrees of freedom across sets, as illustrated for previous designs.

Basic models and analyses of variance of triallel and quadrallel designs repeated over environments are the same as those given for the diallel in Table 4.16 for model II. Direct \( F \)-tests of crosses and crosses with environment mean squares are available to determine if further partitioning of the crosses sums of squares is informative. If the crosses mean square is non-significant, further analyses will be fruitless. For diallel analyses the crosses sums of squares were orthogonally partitioned into average performance of a parent (GCA) and interaction of parents in specific crosses...
Table 4.28 Number of possible hybrids available for different numbers of parents, ignoring reciprocal crosses

<table>
<thead>
<tr>
<th>Number of hybrids</th>
<th>Single crosses</th>
<th>Three-way crosses</th>
<th>Double crosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of parents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>60</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>168</td>
<td>210</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>252</td>
<td>378</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>360</td>
<td>630</td>
</tr>
<tr>
<td>20</td>
<td>190</td>
<td>3,420</td>
<td>14,535</td>
</tr>
<tr>
<td>50</td>
<td>1,225</td>
<td>58,800</td>
<td>690,900</td>
</tr>
</tbody>
</table>

\[ N = \frac{n(n-1)}{2} \times \frac{n(n-1)(n-2)(n-3)}{8} \]

(SCA). Orthogonal partitioning of the triallel and the quadrallel also depends on the number of lines common among crosses and the arrangement of lines within crosses. For example, the \(12\left[4 \times 3 \times 2\right]/2\) three-way crosses possible from four parent lines (A, B, C, and D) are as follows:

1. \((A \times B)C\)  
2. \((A \times B)D\)  
3. \((A \times C)B\)  
4. \((A \times C)D\)  
5. \((A \times D)B\)  
6. \((A \times D)C\)  
7. \((B \times C)A\)  
8. \((B \times C)D\)  
9. \((B \times D)A\)  
10. \((B \times D)C\)  
11. \((C \times D)A\)  
12. \((C \times D)B\)

From the basic model of the triallel, \(Y_{i(jk)} = u + r_i + C_{i(jk)} + e_{i(jk)}\), we can partition the \(C_{i(jk)}\) cross sums of squares as shown by Rawlings and Cockerham (1962a) (Table 4.29), where \(C_{i(jk)}\) is defined as a linear function of uncorrelated effects as \(C_{i(jk)} = (g_i + g_j + g_k) + (s_{2ij} + s_{2ik} + s_{2jk}) + s_{3ijk} + o_{1i} + o_{1(j)} + o_{1(k)} + (o_{2aij} + o_{2aik} + o_{2ak}) + (o_{2bi(j)} + o_{2bi(k)}) + o_{3ijk}\). From the linear model, expected mean squares in Table 4.29 are expressed in terms of components of variance. The components of variance have the following interpretations:

\(\hat{\sigma}_g^2\) = average effect of lines averaged over all orders, e.g., 1, 2, 3, 4, 5, 6, 7, 9, and 11 for A

\(\hat{\sigma}_{s_{2}}^2\) = two-line interaction effect of lines appearing together averaged over all orders, e.g., 1, 2, 3, 5, 7, and 9 for \(A \times B\)

\(\hat{\sigma}_{s_{3}}^2\) = three-line interaction effect of lines appearing together averaged over all orders, e.g., 1 for \((A \times B)C\)

\(\hat{\sigma}_{o_{1i}}^2\) = one-line order effect of lines as a parent, e.g., 7, 9, and 11 for A

\(\hat{\sigma}_{o_{2a}}^2\) = two-line order interaction effects of lines averaged over orders, e.g., 3, 5, 7, and 9 for A and B

\(\hat{\sigma}_{o_{2b}}^2\) = two-line order interaction effects of parent and grandparent lines due to particular order, e.g., 7 and 9 for A and B

\(\hat{\sigma}_{o_{3}}^2\) = three-line order interaction effects of parent and grandparent lines due to particular order, e.g., 1 for \((A \times B)C\)
Table 4.29  Sources of variation for the triallel analysis of variance

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three-way crosses</td>
<td>3(pC_3-1)</td>
<td>(C^*)</td>
<td></td>
</tr>
<tr>
<td>One-line general</td>
<td>(p_1)</td>
<td>(G^*) (\hat{\sigma}^2 + 3r\hat{\sigma}<em>{S_3}^2 + 6rp_3\hat{\sigma}</em>{S_2}^2 + (3rp_2p_3/2)\hat{\sigma}_{a}^2)</td>
<td></td>
</tr>
<tr>
<td>Two-line specific</td>
<td>(pp_3/2)</td>
<td>(S_2^*) (\hat{\sigma}^2 + 3r\hat{\sigma}<em>{S_3}^2 + 3rp_4\hat{\sigma}</em>{S_2}^2)</td>
<td></td>
</tr>
<tr>
<td>Three-line specific</td>
<td>(pp_1p_5/6)</td>
<td>(S_3^*) (\hat{\sigma}^2 + 3r\hat{\sigma}_{S_3}^2)</td>
<td></td>
</tr>
<tr>
<td>One-line order</td>
<td>(p_1)</td>
<td>0(<em>1^*) (\hat{\sigma}^2 + r\hat{\sigma}</em>{0}\hat{\sigma}<em>{a}^2 + 3rp_3\hat{\sigma}</em>{0b}^2 + (rp_2/3)\hat{\sigma}_{0a}^2)</td>
<td></td>
</tr>
<tr>
<td>Two-line order (a)</td>
<td>(pp_3/2)</td>
<td>0(<em>{2a}^*) (\hat{\sigma}^2 + r\hat{\sigma}</em>{0}\hat{\sigma}<em>{a}^2 + (2rp_1/3)\hat{\sigma}</em>{2a}^2)</td>
<td></td>
</tr>
<tr>
<td>Two-line order (b)</td>
<td>(p_1p_2/2)</td>
<td>0(<em>{2b}^*) (\hat{\sigma}^2 + r\hat{\sigma}</em>{0}\hat{\sigma}<em>{a}^2 + 2rp_3\hat{\sigma}</em>{0b}^2)</td>
<td></td>
</tr>
<tr>
<td>Three-line order</td>
<td>(pp_2p_4/3)</td>
<td>0(<em>3^*) (\hat{\sigma}^2 + r\hat{\sigma}</em>{0}\hat{\sigma}_{a}^2)</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>((r-1))(3(pC_3-1))</td>
<td>(E^*) (\hat{\sigma}^2)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Rawlings and Cockerham (1962a)

General (\(g\)) and specific (\(s_2, s_3\)) effects are analogous to effects of diallel analyses, but adding order effects to former analyses. The order effects (\(o_1, o_{2a}, o_{2b}, o_3\)) occur because of the arrangement of parents and line ancestry in three-way crosses. F-tests to determine which effects are significantly different from zero can be made directly for all except one-line order effects; similar to previous mating designs, comparisons of mean squares determine if differences exist but do not provide any genetic information. Again, it becomes necessary to express E(MS) in terms of covariances of relatives because their composition can be expressed in terms of genetic components of variance.

Rawlings and Cockerham (1962a) showed that there were nine covariances of relatives of three-way cross relatives that could be translated into genetic components of variance, whereas in previous mating designs we usually had only two covariances of relatives. The triallel analysis has seven mean squares for estimation of components of variance. Translating from covariances of relatives, genetic components for components of variance in Table 4.29 are shown in Table 4.30.

From Table 4.30, components of variance \(\hat{\sigma}^2_{g}\) and \(\hat{\sigma}^2_{01}\) include only additive effects and additive \(\times\) additive epistatic effects. Variance components \(\hat{\sigma}^2_{S_2}\) and \(\hat{\sigma}^2_{02a}\) include dominance and all types of epistasis or deviations from an all-additive model; \(\hat{\sigma}^2_{S_3}\) includes all types of epistatic effects except additive \(\times\) additive.

If normality of genetic effects in the model is assumed, F-tests can be used to test some genetic hypotheses; e.g., the null hypothesis that \(\hat{\sigma}^2_{g} = 0 = G/S_2\) will be a test for the relative importance of additive genetic effects. If tests of hypotheses indicate significant genetic effects, it would be logical to proceed to obtain estimates of \(\hat{\sigma}^2_{A}, \hat{\sigma}^2_{D}, \hat{\sigma}^2_{AA}, \) etc.
The quadrallel model and analysis are similar in form to the triallel, but covariances of relatives and coefficients of genetic components of variance are different. There are eight covariances of relatives and seven orthogonal partitions of double crosses sums of squares. Tests of hypothesis of appropriate mean squares and their genetic interpretations are similar to those of the triallel. Those interested in the quadrallel analysis should refer to Rawlings and Cockerham (1962b).

Triallel and quadrallel mating designs and analyses are of interest because of the number of independent mean squares for estimation of genetic components of variance. With most of the mating designs the assumption of no epistasis was imposed to provide estimates of $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$. Because we have more mean squares available, it is possible to estimate some of the lower order epistatic variances. If one writes the genetic expectations of mean squares (and there are eight mean squares, including one for experimental error, in the triallel and the quadrallel) in terms of genetic components of variance, the equations may be solved by least squares, weighted least squares, or maximum likelihood to obtain estimates of the genetic model tested. The first models tested may exclude epistasis to determine the degree of fit. If deviations are significant, a model that includes lower order epistatic terms (e.g., $\hat{\sigma}^2_{AA}$) may be fit. Because the mean squares have different variances, Hayman (1958) suggested the method of maximum likelihood to solve the equations. Because one iteration of the maximum likelihood method is often sufficient, the method of weighted least squares may be preferable for estimation.

Triallel and quadrallel mating designs have not been used extensively because they (1) are relatively new in development, (2) are complex mating designs with complex analyses, (3) require a large number of crosses to sample the population adequately, and (4) require two or more growing seasons to produce crosses before they can be tested. Because both analyses are for a specific reference population, most breeding programs do not have an unselected group of parent lines from one
Inbred Lines

4.11 Inbred Lines

Use of inbred lines as parents for diallel, design II, triallel, and quadrallel mating designs has been discussed. In all instances the crosses tested were non-inbred, but population. The large number of crosses to test can be reduced by grouping parent lines into sets. Table 4.28 shows that the number of crosses can be drastically reduced for a group of 60 parent lines, e.g., for the triallel crosses the numbers are reduced from 102,660 for one set of 60 parent lines to 600 for 10 sets of 6 lines in each set. Wright et al. (1971) used diallel and triallel mating designs to estimate genetic variances in a synthetic, Krug Hi I Synthetic 3. Mean squares from the diallel and the triallel were expressed as functions of genetic components of variance and used to test genetic models that included $\hat{\sigma}_A^2$, $\hat{\sigma}_D^2$, and digenic epistasis. Epistasis was not estimable from this analysis, even though 11 equations were available to fit $\hat{\sigma}_A^2$, $\hat{\sigma}_D^2$, $\hat{\sigma}_{AA}^2$, and experimental error. It seems the potential of the triallel and the quadrallel for estimation of genetic variances in a reference population is limited also because of complexity in obtaining the parents and crosses. Generally, such selected lines are of diverse origins, and the estimates of genetic variances would not be valid. As for diallel analyses, effects may be estimated for selected lines, but these estimations apply only to the lines with which they were crossed; the effects could and probably will have different effects when crossed to another group of lines.

Formulation of the composition of genetic variances among single crosses, three-way crosses, and double crosses from a common group of parent lines, however, has an important significance in maize breeding programs that are committed to inbred line and hybrid development. Cockerham (1961) showed that variation among single crosses is always greater than among three-way crosses, and variation among three-way crosses is always greater than among double crosses. If we assume all parents have an inbreeding coefficient of $F = 1$, the genetic variance content of the crosses components of variance (Table 4.28) is shown in Table 4.31.

Table 4.31  Coefficients of components of genetic variance of $\hat{\sigma}_C^2$ among unrelated single, three-way, and double crosses

<table>
<thead>
<tr>
<th>Type of crosses</th>
<th>$\hat{\sigma}_A^2$</th>
<th>$\hat{\sigma}_D^2$</th>
<th>$\hat{\sigma}_{AA}^2$</th>
<th>$\hat{\sigma}_{AD}^2$</th>
<th>$\hat{\sigma}_{DD}^2$</th>
<th>$\hat{\sigma}_{AAA}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Three-way</td>
<td>$\frac{3}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{9}{16}$</td>
<td>$\frac{3}{8}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{27}{64}$</td>
</tr>
<tr>
<td>Double</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{8}$</td>
<td>$\frac{1}{16}$</td>
<td>$\frac{1}{8}$</td>
</tr>
</tbody>
</table>

If only additive genetic effects are assumed, the relative advantage of single, three-way, and double crosses is $1 : \frac{3}{4} : \frac{1}{2}$; i.e., variation among single crosses will be twice that of double crosses if we have only additive effects. If non-additive variance is important, the relative advantage increases for single crosses over three-way and double crosses.

4.11 Inbred Lines

Use of inbred lines as parents for diallel, design II, triallel, and quadrallel mating designs has been discussed. In all instances the crosses tested were non-inbred, but
Hereditary Variance: Mating Designs

The parents could be non-inbred \((F = 0)\), partially inbred \((0 < F < 1)\), or completely inbred \((F = 1)\). The effect of using inbred parents was to increase coefficients of components of genetic variance for covariances of relatives.

For \(F = 0\), \(\text{Cov HS} = (\frac{1}{4}) \hat{\sigma}^2_A + (\frac{1}{16}) \hat{\sigma}^2_{AA}\)

and, for \(F = 1\), \(\text{Cov HS} = (\frac{1}{2}) \hat{\sigma}^2_A + (\frac{1}{4}) \hat{\sigma}^2_{AA}\).

Another method for estimation of genetic variances in a population is to test the unselected inbred lines themselves. Although no mating design is used, variability among inbred lines can be used as an estimate of genetic variability of a reference population. Inbred lines will refer to all lines that have some measure of inbreeding, such as \(S_1\) \((F = 0.5)\), \(S_2\) \((F = 0.75)\), and lines assumed to be completely inbred \((F = 1)\).

Use of inbred lines for estimation of genetic variances requires the same assumption that was imposed when inbred lines were used as parents in crosses; i.e., the inbred lines are an unselected sample of genotypes from the reference population. Adequate sampling of the reference population is necessary, and no selection is imposed on the \(S_0\) plants included to produce the inbred lines. If \(S_1\) lines are used, adequate sampling with a minimum of natural and artificial selection is not too serious. As the inbreeding is continued, however, deleterious recessives are uncovered and it becomes more difficult to obtain a representative sample of the original \(S_0\) plants of the reference population. The degree of difficulty in maintaining a line from each \(S_0\) plant will depend on the genetic load of the reference population.

The frequency of deleterious recessives in an open-pollinated variety is probably greater than for a synthetic variety formed from elite inbred lines because many of the deleterious recessives probably were purged in the development of elite lines used to form the synthetic. Therefore, except for mutation the frequency of deleterious recessives would be expected to be less in the synthetic populations. Results from research in developing inbred lines from open-pollinated varieties (Eberhart et al., 1966) and from a synthetic variety (Hallauer and Sears, 1973) seem to support this hypothesis. Single-seed descent seems to be the logical method for developing inbred lines beyond the \(S_1\) generation (Brim, 1966). The ideal situation would be to have an inbred line for each \(S_0\) plant sampled in the reference population.

The inbred lines developed are themselves tested to determine variability among lines. If an adequate sample, say 196 lines, is available, the lines may be grown in one large replication or partitioned into sets (e.g., augmented design, sets in reps, and/or reps in sets), as illustrated for the mating designs. A \(14 \times 14\) lattice design also may be an appropriate field design to use for increasing the local control of experimental error. Measurements are made on inbred plots in each environment, and analyses on plot means are the ordinary analysis of variance for entries tested in different environments (Table 4.32). Direct \(F\)-tests and estimates of components of variance can be made from mean squares shown in Table 4.32.

Because the inbred lines are assumed to be a random sample of genotypes from the reference population, model II analyses are appropriate. To obtain estimates of
### Table 4.32 Analysis of variance of inbred lines repeated over environments

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environments (E)</td>
<td>$e-1^a$</td>
<td>$\sigma^2 + re\sigma^2_{ge} + re\sigma^2_g$</td>
<td></td>
</tr>
<tr>
<td>Replications/E</td>
<td>$e(r-1)$</td>
<td>$M_4$</td>
<td></td>
</tr>
<tr>
<td>Inbred lines</td>
<td>$n-1$</td>
<td>$M_1 \quad M_2$</td>
<td></td>
</tr>
<tr>
<td>E $\times$ inbred lines</td>
<td>$(e-1)(n-1)$</td>
<td>$M_3$</td>
<td></td>
</tr>
<tr>
<td>Pooled error</td>
<td>$e(r-1)(n-1)$</td>
<td>$M_2$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$ern-1$</td>
<td>$M_1$</td>
<td></td>
</tr>
</tbody>
</table>

$a, r, n,$ and $k$ refer to the number of environments, replications within environments, inbred lines, and individual plants measured within each plot, respectively.

$b \hat{\sigma}^2 = (\hat{\sigma}^2_{we} + \hat{\sigma}^2_{wg}) / k + \hat{\sigma}^2_p,$ where $\hat{\sigma}^2_p$ is experimental plot error.

---

Genetic variance in the reference population, it is necessary to translate (as shown for previous mating designs) components of variance to genetic components of variance. The genetic composition of the variance component due to inbred lines, for different levels of inbreeding, is shown in Chapter 2. Dominance effects dissipate rapidly with inbreeding. Because we have only one equation, the assumption of no dominance can be made, and the inbred line component of variance provides an estimate of the additive genetic variance. For instance, variance among $S_1$ lines provides an estimate of $\hat{\sigma}^2_A$. It is also shown that when $F = 1$ for the inbred lines, the variation is doubled among those with no genetic variation within the lines. For the case of $F = 1$ the inbred line variance component gives an estimate of $2\hat{\sigma}^2_A$. Depending on the level of inbreeding, therefore, the estimate of $\hat{\sigma}^2_A$ from the inbred line component of variance can range from $\hat{\sigma}^2_A (F = 0.5)$ to $2\hat{\sigma}^2_A (F = 1.0)$. The seriousness of the dominance bias in the estimate of $\hat{\sigma}^2_A$ is shown in Table 2.13. If $p = q = 0.5$ (no dominance effects), variance among lines will be due only to additive effects.

Because estimates of $\hat{\sigma}^2_A$ can be obtained from the inbred line tests, estimates of heritability on a progeny mean basis can be obtained. If, for instance, $S_1$ lines are evaluated in experiments repeated over environments (Table 4.32), an estimate of heritability on $S_1$ line means can be obtained as

$$\hat{h}^2 = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2 / re + \hat{\sigma}^2_{ge} / e + \hat{\sigma}^2_g}$$

Because $S_1$ lines are used, the bias from dominance effects is $(\frac{1}{4})\hat{\sigma}^2_D$ if $p \neq q$ (dominance effects are present). The estimate of heritability is not biased by genotype–environment interaction effects because an estimate of $\hat{\sigma}^2_{AE}$ is provided by the combined analysis. Approximate standard error of the heritability estimate is

$$\text{SE}(\hat{h}^2) = \frac{\text{SE}(\hat{\sigma}^2_g)}{\hat{\sigma}^2 / re + \hat{\sigma}^2_{ge} / e + \hat{\sigma}^2_g}$$
where the standard error of $\hat{\sigma}_g^2$ is obtained from Table 4.32 as the square root of

$$\frac{2}{(re)^2} \left[ \frac{M_4^2}{n+1} + \frac{M_3^2}{(e-1)(n-1)+2} \right]$$

Because the component of variance $\hat{\sigma}_g^2$ estimates $\hat{\sigma}_A^2$, we do not have a coefficient for the standard error of $\hat{\sigma}_A^2$, as in previous mating designs.

If lines that are near homozygosity are evaluated, we do not have any bias due to dominance included in our estimate of $\hat{\sigma}_A^2$ from the component $\hat{\sigma}_g^2$. For $F = 1$, $\hat{\sigma}_g^2 = 2\hat{\sigma}_A^2$, and an estimate of heritability based only on additive effects can be obtained in the usual way.

Use of inbred lines for estimation of $\hat{\sigma}_A^2$ of a reference population seems enticing, but use of inbred lines has at least two serious handicaps. First, one needs to develop a set of unselected inbred lines that are representative of genotypes of the reference population; this becomes increasingly difficult as the level of inbreeding increases. Second, the time required to develop inbred lines, particularly for greater levels of inbreeding, would be greater than for mating designs using non-inbred parents unless doubled haploids are utilized (see Chapter 1). Neither handicap is present if $S_1$ inbred lines are used. There has been some speculation that information obtained from inbred line tests would not be as good as information obtained from use of non-inbred material because of larger experimental errors associated with inbred line tests. Present evidence, however, does not indicate that experimental errors of inbred tests are any greater than those for non-inbred tests. Use of $S_1$ lines seems to be a good method for estimation of $\hat{\sigma}_A^2$ in maize populations if departures from certain assumptions (e.g., $p = q = 0.5$ and no dominance) are not serious.

### 4.12 Selection Experiments

All mating designs discussed require that progenies be developed and their evaluation in experiments repeated over environments to provide estimates of genetic components of variance. To ensure proper sampling of the reference population, large experiments are needed to estimate genetic parameters with reasonable sampling errors. Also, experiments need to be repeated over environments to obtain estimates of genetic parameters unbiased by environmental effects. Consequently, conducting experiments to obtain genetic information can be very expensive in time and resources. The primary purpose for obtaining estimates of genetic parameters is to provide guidelines in developing breeding programs based on populations with diverse genetic structures (e.g., target-specific breeding methods for specific populations) and to predict future gain from selection. Often the breeder has populations under selection for different traits but has not determined relative proportions of differences among progenies due to genetic and to environmental forces. Selection experiments, which include adequate sampling and testing in properly conducted
field experiments repeated over environments, also provide information necessary for predicting future gain from selection. One case of a selection experiment not providing genetic information is mass selection, where we do not have progenies for testing in replicated tests.

Any selection methods that use progeny information for selecting the best individuals to recombine for population improvement will provide estimates of genetic variance (under the assumption of no non-additive effects), genetic–environmental interaction variance, and experimental error. If we use Table 4.32 as an example, direct estimates of these components can be determined from linear functions of the mean squares. Genetic composition of the estimate of $\hat{\sigma}_g^2$ depends on the type of progeny evaluation. Table 1.2 (Chapter 1) lists types of progenies that may be evaluated. For example, if half-sib progenies are evaluated, $\hat{\sigma}_g^2$ contains $(\hat{\nu}_g)\hat{\sigma}_A^2$, and, assuming $F = 0$ and no epistatic variance, $\hat{\sigma}_A^2 = 4\hat{\sigma}_g^2$. If $S_1$ progenies are evaluated, $\hat{\sigma}_g^2 = \hat{\sigma}_A^2$, assuming $p = q = 0.5$ or no dominance effects. Since $\hat{\sigma}_G^2$ is estimated as $M_4 - M_3$, the variance of our estimate is

$$[2/(r^2e^2)][M_4^2/((n + 1)M_3^2/((e - 1)(n - 1)] + 2$$

Similarly, estimates of the interaction of genetic effects with environments and its variance can be calculated from $M_3$ and $M_2$ mean squares. A coefficient of genetic variation can be calculated as $((\hat{\sigma}_g/\bar{X}) \times 100)$, where $\bar{X}$ is the mean of all progenies and estimates of heritability can be calculated as shown for the case of inbred lines.

For the initial sampling of a population undergoing selection, estimates of $\hat{\sigma}_A^2$ should be valid if experiments are repeated over environments, assuming no epistasis. Although estimates of $\hat{\sigma}_A^2$ may be biased upward by dominance and epistatic effects, the bias will not be any more serious than for the assumption of no epistasis imposed on the analysis for most of the mating designs. We could speculate epistatic effects could be zero on average or included in other genetic components of variance. Additive genetic variance is the component of genetic variance useful to the breeder in a selection program because it is fixable; non-additive intraloci effects (dominance) and interloci effects involving dominance (epistasis) are not fixable. Estimates of $\hat{\sigma}_A^2$ are dependent on gene frequency. If selection is effective, which is the primary objective of selection (i.e., changing gene frequency of the trait(s) toward a desirable direction), estimates of $\hat{\sigma}_A^2$ will change in cyclical selection programs. An estimate of $\hat{\sigma}_A^2$ however, is available from each cycle of selection. Estimates of $\hat{\sigma}_A^2$ may change with cycles of selection but should be equally valid if sampling was sufficient (which is necessary for effective selection) for each cycle and extensive testing was conducted for progeny evaluation to separate environmental from genetic effects. Estimates of $\hat{\sigma}_A^2$ will vary among cycles (because of sampling) and introduce an unknown source of error (e.g., molecular markers targeted at selection would need to be re-evaluated or ‘re-trained’ at each selection cycle). Standard errors for estimates of $\hat{\sigma}_A^2$ can be calculated to determine confidence limits in successive cycles of selection. Changes in gene frequency for most
quantitative traits are not expected to be very large for several cycles of selection. Thus estimates of $\hat{\sigma}_A^2$ for a few selection cycles may be similar but not equal due to sampling. A pooled estimate of $\hat{\sigma}_A^2$ can be obtained from the pooled analysis that could be an excellent parameter for predicting selection for the next three to five cycles. However, variation in estimates of $\hat{\sigma}_A^2$ for different cycles of selection also can be influenced by the extent of recombination of selected parents to form the next cycle population. Effects of recombination could be particularly evident if selected parents have more generations of inbreeding (e.g., $S_2$) because of retention of large linkage blocks. Environmental effects for different cycles could be important in expansion or contraction of variability among progenies tested (e.g., the range among progenies may be expanded in a favorable environment and compressed in a very unfavorable one). Pooled estimates of $\hat{\sigma}_A^2$ would tend to dilute the extremes and give a more valid estimate when pooled or combined across environments within cycles and across cycles. The genetic coefficient of variation calculated for each cycle also would provide a measure of the range among progenies relative to the overall mean for progenies.

Estimates of genetic variance from selection experiments that use progeny evaluation is predicated on the thesis that the researcher is initiating a long-term cyclical selection program. If the decision is made to initiate such a program, estimates of $\hat{\sigma}_A^2$ calculated from $\hat{\sigma}_g^2$ to predict expected progress are better than those obtained from special mating designs. Also, research activity can be given to the selection study without diverting resources to a special study. If the experimenter, however, desires to have estimates for a specific population to answer a specific question, the only recourse is to use a specific mating design. It must be emphasized that adequate sampling and evaluation are as important for valid estimates of genetic parameters as for a long-term selection experiment that can involve the span of the researcher’s career (Lamkey and Hallauer, 1987). Marker-based breeding schemes are an alternative now that genotyping cost is reduced. However, these strategies should be focused. Finding QTL for complex traits has been easy but exploiting them in selection has demonstrated to be challenging in the past 20 years (Bernardo, 2008) and marker-based selection without QTL mapping has been proposed and remains to be validated, e.g., genome-wide selection in maize (Bernardo and Yu, 2007).

4.13 More on $F_2$ Populations (Special Case of $p = q = 0.5$)

Gene frequencies of populations described so far generally are not known. It is shown in Chapter 2 that means and genetic variances of populations depend on gene frequencies and change as the frequencies change. In certain types of populations, however, expected average gene frequency is known; they are $F_2$ populations produced from inbred lines (equivalent to pure lines or cultivars in self-pollinated species) that are assumed to be homozygous and homogeneous ($F = 1$). If we cross two inbred lines to form the $F_1$ hybrid and then self-pollinate the heterozygous but homogeneous $F_1$ plants, we derive the $F_2$ population (heterozygous and heterogeneous) that will segregate at each heterozygous locus in the
4.13 More on \( F_2 \) Populations (Special Case of \( p = q = 0.5 \))

Table 4.33 Distribution of genotypes and genotypic values of an \( F_2 \) population formed by crossing two inbred lines, where \( A \) is the desirable allele and \( a \) is the undesirable allele

<table>
<thead>
<tr>
<th>Genotypic</th>
<th>Frequency of genotype</th>
<th>Genotypic value</th>
<th>Coded genotypic value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>( p^2 = 0.25 )</td>
<td>( C + a )</td>
<td>( a )</td>
</tr>
<tr>
<td>Aa</td>
<td>( 2pq = 0.50 )</td>
<td>( C + d )</td>
<td>( d )</td>
</tr>
<tr>
<td>Aa</td>
<td>( q^2 = 0.25 )</td>
<td>( C - a )</td>
<td>( -a )</td>
</tr>
</tbody>
</table>

well-known Mendelian ratio of \( \frac{1}{4} \) homozygous dominant, \( \frac{1}{2} \) heterozygous, and \( \frac{1}{4} \) homozygous recessive (Table 4.33).

If the two parents have the same allele at some loci, there will be no segregation in the \( F_2 \) population. Because we cross two inbred lines, average gene frequency at segregating loci will be 0.5; hence, expected frequencies of genotypes are known. The constant \( C \) for genotypic values in Table 4.33 includes actions of all genes not under consideration as well as non-heritable forces. Mather (1949) and Mather and Jinks (1971) have used the model given in Table 4.33 to develop expected means and genetic components of variance for the different types of generations that can be produced from crossing two inbred lines. The average gene frequency of 0.5 can be viewed from two standpoints: (1) average gene frequency of desirable alleles and undesirable alleles will be 0.5 at each locus or (2) average gene frequency of desirable alleles and undesirable alleles will be 0.5 over all loci.

For estimation of genetic components of variance, linkages of genes, and interloci genetic effects (which in most if not all instances are not valid) are assumed absent. Linkage effects on estimates of genetic components of variance can be reduced by random mating of the \( F_2 \) population until linkage equilibrium of the genes is approached. We will always have linkages, but an equilibrium of linkages can be approached by random mating. Epistatic effects probably are also present, but Mather (1949) has devised tests to determine seriousness of epistatic effects. Tests for non-additivity of genetic effects among loci include comparing different generations and how much they deviate from the additive model. If non-additive effects are detected among loci, Mather has suggested that a transformation of the data be made before analysis. Transformation to an additive scale among loci would satisfy the assumption of no epistasis for estimation of main effect – additive and dominance variance (non-additivity within loci), which is total genetic variance of the two alleles at each locus. From Table 4.33 we have at each locus three genotypes or 2 df that can be partitioned as (1) additive genetic variance, which is the sum of squares due to regression; and (2) dominance variance, due to deviations from regression.

Mather (1949) and Mather and Jinks (1971) developed procedures to arrive at the means and the proportions of total genetic variance due to additive and dominance effects. For the special case of \( p = q = 0.5 \), \( F_2 \) is our reference population and is equivalent to the open-pollinated, synthetic, and composite variety populations used as reference populations for other mating designs for interpreting genetic
components of variance. Hence the $S_0$ plants of a variety are equivalent to $F_2$ plants, and the populations are similar in genetic structure (see Table 2.2 vs. Table 4.33) for developing the relations. The two inbred parents and the $F_1$ hybrid of the two parents contain no genetic variability and are merely building blocks for constructing the $F_2$ population. From Table 4.33 and procedures outlined in Chapter 2 the mean of the $F_2$ population is

$$\left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D$$

for one locus

Similarly, total genetic variance in the $F_2$ population is

$$\left(\frac{1}{4}\right)a^2 + \left(\frac{1}{2}\right)d^2 + \left(\frac{1}{4}\right)a^2 - \left[\left(\frac{1}{2}\right)d\right]^2 = \left(\frac{1}{2}\right)a^2 + \left(\frac{1}{4}\right)d^2$$

for one locus

Summing over all independent loci segregating for the trait in question, the mean is $\left(\frac{1}{2}\right)\sum_i d_i$ and total genetic variance is $\left(\frac{1}{2}\right)\sum_i a_i^2 + \left(\frac{1}{4}\right)\sum_i d_i^2$. For brevity we will designate the mean as $\left(\frac{1}{2}\right)d$ and total genetic variance as $\left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D$.

Note: For consistency we will use $A$ for Mather’s $D$ and $D$ for Mather’s $H$ to conform with the usage of $\hat{\sigma}^2_A$ for additive genetic variance and $\hat{\sigma}^2_D$ for dominance variance.

Relations shown in Chapter 2 and the ones developed for the $F_2$ population can be shown to be equivalent if we refer to the basic definitions of the mean and total genetic variance of a population. It is shown in Chapter 2 that the mean is $(p - q)a + 2pqd$, which on substitution of $p = q = \frac{1}{2}$ is $\left(\frac{1}{2}\right)d$

Total genetic variance can be represented by:

$$2pq(a + (q - p)d) + 4p^2 q^2 d^2,$$

which on substitution is $\left(\frac{1}{2}\right)a^2 + \left(\frac{1}{4}\right)d^2$

Hence the relations of the total genetic variance of a reference population: $\hat{\sigma}^2_A + \hat{\sigma}^2_D$ and $\left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D$ are equivalent except that the expected gene frequency has been substituted in the definitions of $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$ in the latter.

Total genetic variance of the $F_2$ population can be estimated by use of the parents and/or the $F_1$ generation (to provide an estimate of the environmental effects) and the $F_2$ population. If we desire to determine relative proportions of total $F_2$ genetic variance due to additive and dominance effects, we need to develop additional generations. One common procedure is to self-pollinate individual $S_0$ plants in the $F_2$ generation to advance to the $S_1$ level, commonly referred to as the $F_3$ generation. For maize, controlled self-pollinations are needed. If individual plants in the $F_2$ generation are measured, parent–offspring regressions also can be determined by regressing the $F_3$ progeny means (values of $S_1$) on the $F_2$ parent plants (values of $S_0$), which is

$$\hat{\sigma}^2_A + \left(\frac{1}{2}\right)\hat{\sigma}^2_D$$

or $\left(\frac{1}{2}\right)A + \left(\frac{1}{8}\right)H$ in Mather’s (1949) notation

Again, it is important to adequately sample the $F_2$ population for estimates of genetic components of variance to be representative of the population variability. The genotypic array of the $S_1$ ($=F_3$) generation is

$$(\left(\frac{1}{4}\right)AA + \left(\frac{1}{2}\right)[(\left(\frac{1}{4}\right)AA + \left(\frac{1}{2}\right)Aa + \left(\frac{1}{4}\right)aa] + \left(\frac{1}{4}\right)aa$$

which, by collecting frequencies of different classes of genotypes, becomes
The $S_1$ generation mean is determined as $\left(\frac{3}{8}\right)a + \left(\frac{1}{4}\right)d + \left(\frac{3}{8}\right)(-a) \text{ or } \left(\frac{1}{4}\right)d$

Two sources of genetic variation in the $F_3$ generation are

1. variation among $F_3$ progeny means and
2. mean variation of $F_3$ progenies.

Means of the $F_3$ progeny array are $\left(\frac{1}{4}\right)a$ for AA, $\left(\frac{1}{2}\right)d$ for Aa, and $\left(\frac{1}{4}\right)(-a)$ for aa. Hence the $F_3$ progeny variance $\hat{\sigma}^2_{gF_3}$ is determined as

\[
\hat{\sigma}^2_{gF_3} = \left(\frac{1}{4}\right)a^2 + \frac{1}{2}d^2 + \left(\frac{1}{4}\right)(-a)^2 - \left(\frac{1}{2}\right)d^2 = \left(\frac{1}{2}\right)a^2 + \frac{1}{8}d^2 - \frac{1}{16}d^2 = \left(\frac{1}{2}\right)a^2 + \frac{1}{16}d^2
\]

By summing over loci we have $\hat{\sigma}^2_{gF_3} = \left(\frac{1}{2}\right)a + \left(\frac{1}{16}\right)d$, which is equal to $\hat{\sigma}^2_{A} + \left(\frac{1}{2}\right)\hat{\sigma}^2_{D}$ from the definitions in Chapter 2.

The mean variance of $F_3$ progenies ($\hat{\sigma}^2_{F_3}$) is the variation within the progenies. From the genotypic arrays, the AA and aa types will not contain any genetic variation and need not be considered further. The $F_3$ progenies originating from the Aa $F_2$ plants will have segregation ratios similar to the $F_2$ population. The frequency of this progeny in the $F_3$ generation is 0.5 and thus equals $\left(\frac{1}{2}\right)\left[\left(\frac{1}{4}\right)AA + \left(\frac{1}{2}\right)Aa + \left(\frac{1}{4}\right)aa\right]$ with a mean of $\left(\frac{1}{4}\right)d$. The mean $F_3$ variance ($\hat{\sigma}^2_{F_3}$) becomes

\[
\hat{\sigma}^2_{F_3} = \left(\frac{1}{2}\right)a^2 + \frac{1}{4}d^2 - \frac{1}{16}d^2 = \left(\frac{1}{2}\right)a^2 + \frac{1}{8}d^2
\]

Summing over all loci, $\hat{\sigma}^2_{F_3} = \left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D$, which is equal to $\left(\frac{1}{2}\right)\hat{\sigma}^2_{A} + \left(\frac{1}{2}\right)\hat{\sigma}^2_{D}$.

If we want to sum the effects for all loci, effects for individual loci can be added (since we assumed effects for each locus are additive or made so by transformation); and we have the expressions for the phenotypic variances given by Mather (1949) and Mather and Jinks (1971):

\[
\hat{\sigma}^2_{F_2} = \left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D + E_1
\]
\[
\hat{\sigma}^2_{F_3} = \left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D + E_2
\]
\[
\hat{\sigma}^2_{F_3} = \left(\frac{1}{2}\right)A + \left(\frac{1}{4}\right)D + E_1
\]

where the $E$ components are the non-heritable variances associated with the respective components of variance.

The same techniques can be used to derive the mean and variance of subsequent generations ($F_4$, $F_5$, etc.) derived from the $F_2$ population (Mather, 1949). Backcross populations derived from the inbred lines and their $F_1$ populations are similar to $F_2$ populations because each $S_0$ plant is a unique entity and cannot be
repeated. Each $S_0$ plant can be measured, but no estimate of error is available. However, measurements of individual $S_0$ backcross plants can be used to compute parent–offspring regressions.

Backcross generations can be advanced by selfing individual $S_0$ backcross plants, as was done by advancing $F_2$ plants to the $F_3$ generation. If the $F_1$ hybrid $Aa$ is backcrossed to the parent having the desirable allele, the backcrossed genotypic array ($BC_1$) becomes $(\frac{1}{2})AA + (\frac{1}{2})Aa$. By substituting genotypic values for respective genotypes:

the $BC_1$ mean becomes $(\frac{1}{2})a + (\frac{1}{2})d$ and the $BC_1$ variance is

$$(\frac{1}{2})a^2 + (\frac{1}{2})d^2 - [(\frac{1}{2})a + (\frac{1}{2})d]^2 = (\frac{1}{4})A + (\frac{1}{4})D - (\frac{1}{2})AD + E_1$$

by summing over all loci and including the non-heritable variance $E_1$.

If we self individual backcross plants, the genotypic array of the backcrossed-selfed ($BS_1$) generation progenies becomes $(\frac{1}{2})AA + (\frac{1}{4})(\frac{1}{2})AA + (\frac{1}{2})Aa + (\frac{1}{4})aa$. Substituting genotypic values for genotypes, we find that the $BS_1$ mean is $(\frac{1}{2})a + (\frac{1}{4})(\frac{1}{2})a + (\frac{1}{2})d + (\frac{1}{4})(-a)$ or $(\frac{1}{2})a + (\frac{1}{4})d$.

The variance among $BS_1$ means is $\hat{\sigma}^2_{BS_1} = (\frac{1}{2})a^2 + (\frac{1}{2})(\frac{1}{2})d^2 - [(\frac{1}{2})a + (\frac{1}{4})d]^2$, which becomes $(\frac{1}{4})A + (\frac{1}{16})D - (\frac{1}{4})AD + E_2$ after collecting terms and summing over all loci.

Mean variance of $BS_1$ progenies ($\hat{\sigma}^2_{BS_1}$) would involve only the progenies that arose from selfing the $Aa$ plants and is similar to the mean variance in the $F_3$, hence, $\hat{\sigma}^2_{BS_1} = (\frac{1}{4})a^2 + (\frac{1}{8})D - (\frac{1}{4})AD + E_2$.

The same calculations would be needed to determine the variance of the population formed by crossing the $F_1$ hybrid to the parent having the opposite allele, which is designated as the $BC_2$ population. Plants of the $BC_2$ are selfed to produce the $BS_2$ progenies. The $BC_2$ population of $(\frac{1}{2})Aa + (\frac{1}{2})aa$ has a mean of $(\frac{1}{2})d - (\frac{1}{2})a$ and a variance of $(\frac{1}{4})A + (\frac{1}{4})D + (\frac{1}{2})AD + E_1$. Selfed progenies of the $BC_2$ population also contain two sources of variation:

(1) variance among $BS_2$ progenies or $(\frac{1}{4})A + (\frac{1}{16})D - (\frac{1}{4})AD + E_2$ and
(2) mean variance of $BS_2$ progenies or $(\frac{1}{4})A + (\frac{1}{8})D + E_1$.

Expectations of the variances of the two backcross populations can be summarized as follows:

$$\hat{\sigma}^2_{BC_1} = (\frac{1}{4})A + (\frac{1}{4})D - (\frac{1}{2})AD + E_1$$
$$\hat{\sigma}^2_{BC_2} = (\frac{1}{4})A + (\frac{1}{4})D + (\frac{1}{2})AD + E_1$$
$$\hat{\sigma}^2_{BS_1} = (\frac{1}{4})A + (\frac{1}{16})D - (\frac{1}{4})AD + E_2$$
$$\hat{\sigma}^2_{BS_2} = (\frac{1}{4})A + (\frac{1}{16})D + (\frac{1}{4})AD + E_2$$
$$\hat{\sigma}^2_{BS_1} = (\frac{1}{4})A + (\frac{1}{8})D + E_1$$
$$\hat{\sigma}^2_{BS_2} = (\frac{1}{4})A + (\frac{1}{8})D + E_1$$
In the BC₁, BC₂, BS₁, and BS₂ generations the additive and dominance effects are not separable; if they are of similar magnitude, one can combine the equations for estimates of $D$ and $H$.

For instance,

\[
\hat{\sigma}^2_{\text{BC}_1} + \hat{\sigma}^2_{\text{BC}_2} = \left(\frac{1}{2}\right)A + \left(\frac{1}{2}\right)D + 2E_1
\]

and

\[
\hat{\sigma}^2_{\text{BS}_1} + \hat{\sigma}^2_{\text{BS}_2} = \left(\frac{1}{2}\right)A + \left(\frac{1}{8}\right)D + 2E_2
\]

As for $F_2$ and $F_3$ populations, $E_1$ and $E_2$ terms refer to the error variance among individual plants and progeny means, respectively.

Expressions for the different generations can be used to estimate additive and dominance variances. As noted for the expressions, we have two error terms for which we need estimates. An estimate of the error variance for the $F_2$, $BC_1$, and $BC_2$ populations can be obtained as the within variance for environmental effects of the inbred parent lines and their $F_1$ hybrid. Use of environmental effects estimated from inbred parent lines and their $F_1$ hybrids of maize has been subject to criticism because variation among inbred individuals may be greater than expected and that among $F_1$ hybrid individuals less than expected among segregating individuals in an $F_2$ or backcross population. Estimates of $E_2$ obtained by pooling sums of squares and degrees of freedom of inbred parents and $F_1$ hybrids, however, usually are the only estimates available. A technique suggested by Warner (1952) eliminates the necessity of estimating environmental effects on individual plant measurements of inbred parents and their $F_1$ hybrid. Warner’s method provides an estimate of $(\frac{1}{2})A$, which then can be used to calculate an estimate of heritability $h^2$ on an individual plant basis; the method requires measurement of the $F_2$ and both backcross populations:

<table>
<thead>
<tr>
<th>Population</th>
<th>Expected components of variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_2$</td>
<td>$\sigma^2_{F_2} = (\frac{1}{2})A + (\frac{1}{4})D + E_1$</td>
</tr>
<tr>
<td>$2(F_2)$</td>
<td>$2\sigma^2_{F_2} = A + (\frac{1}{2})D + 2E_1$</td>
</tr>
<tr>
<td>$BC_1 + BC_2$</td>
<td>$\sigma^2_{BC_1} + \sigma^2_{BC_2} = (\frac{1}{2})A + (\frac{1}{2})D + 2E_1$</td>
</tr>
</tbody>
</table>

Therefore,

\[
2F_2 - (BC_1 + BC_2) = (\frac{1}{2})A
\]

and heritability in its narrow sense but based on individual $F_2$ plants can be estimated:
\[
\hat{h}^2 = \frac{(\hat{r}_2)A}{\hat{\sigma}^2_{F_2}} = \frac{(\hat{r}_2)A}{(\hat{r}_2)A + (\hat{r}_4)D + E_1}
\]

The method of estimating \((\hat{r}_2)A\) (or \(\hat{\sigma}^2_A\)) does not require a direct estimate of \(E_2\) but it requires the assumption that non-heritable components of variance are comparable for the \(F_2\) and backcross populations. This assumption seems to be logical because the populations (\(F_2\) and combined backcrosses) include the same genotypes, but the combined backcrosses have twice the frequency of the genotypes. A breeder working with a plant species or trait with little information can obtain some preliminary information by this procedure. A relatively high heritability estimate on an individual plant basis would indicate that a simple selection procedure, such as mass selection, would be effective (e.g., flowering time).

Another example of how expected components of variance for different generations can provide genetic information is the evaluation of \(F_3\) progenies, which are, in our nomenclature, equivalent to \(S_1\) lines. Two inbred lines are crossed, their \(F_1\) hybrid is selfed, and an adequate sample of \(F_2\) plants is self-pollinated to produce \(F_3\) progenies, i.e., \(F_3\) seed produced on \(F_2\) plants. If proper pollination techniques are used, adequate \(F_3\) seed should be available for testing in replicated trials. Assume that the experimental entries include the parents (\(P_1\) and \(P_2\)), their \(F_1\) hybrid, and 100 \(F_3\) progenies. One possible form of the analysis of variance of entries included in replicated tests is shown in Table 4.34.

A direct \(F\)-test can be made to determine if the differences among \(F_3\) progenies are significant; if they are

\[
\hat{\sigma}^2_{F_3} = \frac{(M_{32} - M_2)}{r} = (\hat{r}_2)A + (\hat{r}_4)D
\]

which is equivalent to the variation among \(S_1\) progenies, \(\hat{\sigma}^2_A + (\hat{r}_4)\hat{\sigma}^2_D\).

Mean variance of \(F_3\) progenies is estimated from the \(M_{12}\) mean square, which has an expectation of

<table>
<thead>
<tr>
<th>SOV</th>
<th>General</th>
<th>Example</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>(r-1)</td>
<td>3</td>
<td>(M_3)</td>
<td></td>
</tr>
<tr>
<td>Entries</td>
<td>(n-1)</td>
<td>102</td>
<td>(M_{31})</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{\hat{g}})</td>
</tr>
<tr>
<td>Among generations</td>
<td>(g-1)</td>
<td>3</td>
<td>(M_{32})</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{\hat{g}_{F_3}})</td>
</tr>
<tr>
<td>Among (F_3) progenies</td>
<td>(p-1)</td>
<td>99</td>
<td>(M_2)</td>
<td>(\hat{\sigma}^2)</td>
</tr>
<tr>
<td>Error</td>
<td>((r-1)(n-1))</td>
<td>306</td>
<td>(M_{1})</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>(rn-1)</td>
<td></td>
<td>(3708)</td>
<td>(M_{12})</td>
</tr>
<tr>
<td>Within (F_3) entries</td>
<td>(rk(k-1))</td>
<td>3600</td>
<td>(M_{11})</td>
<td>(\hat{\sigma}^2_{\hat{w}})</td>
</tr>
</tbody>
</table>
More on $F_2$ Populations (Special Case of $p = q = 0.5$)

\[ \hat{\sigma}^2_{F_3} = (\nu_4)A + (\nu_6)D + E_1 \]

Because the parents and the $F_1$ hybrid are homogeneous, $M_{11}$ provides an estimate of individual plant environmental effects. An estimate of $E_2$ is provided by the $M_2$ mean square and is an estimate of experimental error on a plot basis for $F_3$ progeny means; i.e., $E_2 = M_2/r$, the variance of a mean. Hence we have four mean squares and four unknowns, A, D, E_1, and E_2:

\[ \hat{\sigma}^2_{F_3} = (\nu_4)A + (\nu_6)D = (M_{32} - M_2)/r \]
\[ \hat{\sigma}^2_{F_3} = (\nu_4)A + (\nu_6)D = M_{12} - M_{11} \]
\[ M_2/r = E_2 \text{ and } M_{11} = E_1 \]

Estimates of $E_1$ and $E_2$ are made directly from the mean squares in Table 4.34. To estimate $A$ and $D$, we have two equations:

\[ (\nu_4)A + (\nu_6)D = (M_{32} - M_2)/r \]
and
\[ (\nu_4)A + (\nu_6)D = M_{12} - M_{11} \]

After solving for $A$ and $D$, we find

\[ A = (\nu_3) \{ [2(M_{32} - M_2)/r] - (M_{12} - M_{11}) \} \]
and,
\[ D = (\nu_6)[2(M_{12} - M_{11}) - (M_{32} - M_2)/r] \]

Variance $V$ of the estimates can be obtained rather easily for $E_1$, $E_2$, and $A$, but as shown in the mating designs, the estimate of $D$ was obtained from a complex function of mean squares:

\[ V_{E_1} = \frac{2M_{11}^2}{r(n-1)+2}, \quad V_{E_2} = \frac{2M_2^2}{(r-1)(n-1)+2} \]
\[ V_A = \frac{128}{9r^2} \left[ \frac{M_{12}^2}{p+1} + \frac{M_2^2}{(r-1)(n-1)+2} \right] + \frac{32}{9} \left[ \frac{M_{12}^2}{rp(k-1)+2} + \frac{M_{11}^2}{rh(k-1)+2} \right] \]
\[ V_D = \frac{512}{9} \left[ \frac{4M_{12}^2}{rp(k-1)+2} + \frac{4M_{11}^2}{r(n-1)+2} + \frac{1}{r} \left[ \frac{M_{32}^2}{p+1} + \frac{M_2^2}{(r-1)(n-1)+2} \right] \right] \]

This example includes the minimum equations necessary to estimate four parameters. We have a perfect fit in this example because the number of equations equals the number of parameters we wished to estimate. If we had included the backcross (BC_1 and BC_2) and backcross-selfed populations, three additional equations would be available for the estimation of $A$ and $D$:
\[ \hat{\sigma}^2_{BS1} + \hat{\sigma}^2_{BS2} = (\hat{\gamma}_2)A + (\hat{\gamma}_8)D + 2E_2 \]
\[ \hat{\sigma}^2_{BS1} = (\hat{\gamma}_4)A + (\hat{\gamma}_8)D + E_1 \]
\[ \hat{\sigma}^2_{BS2} = (\hat{\gamma}_4)A + (\hat{\gamma}_8)D + E_1 \]

Table 4.35 shows an analysis of variance with parents and F\(_1\) (to provide an estimate of environmental effects for individual plants) and the F\(_3\), BS\(_1\), and BS\(_2\) progenies.

Table 4.35 Analysis of variance of an experiment grown in one environment that includes the parents, their F\(_1\) hybrid; and F\(_3\), BS\(_1\), and BS\(_2\) progenies

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r−1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entries</td>
<td>n−1</td>
<td>M(_3)</td>
<td></td>
</tr>
<tr>
<td>Among generations</td>
<td>g−1</td>
<td>M(_{31})</td>
<td></td>
</tr>
<tr>
<td>Entries/groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among F(_3)</td>
<td>p−1</td>
<td>M(_{32})</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{F3})</td>
</tr>
<tr>
<td>Among BS(_1)</td>
<td>s−1</td>
<td>M(_{33})</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{BS1})</td>
</tr>
<tr>
<td>Among BS(_2)</td>
<td>t−1</td>
<td>M(_{34})</td>
<td>(\hat{\sigma}^2 + r\hat{\sigma}^2_{BS2})</td>
</tr>
<tr>
<td>Error</td>
<td>(r−1)(n−1)</td>
<td>M(_2)</td>
<td>(\hat{\sigma}^2)</td>
</tr>
<tr>
<td>Total</td>
<td>rn−1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within F(_3) progenies</td>
<td>rp(k−1)</td>
<td>M(_{11})</td>
<td>(\hat{\sigma}^2_{we} + \hat{\sigma}^2_{wg})</td>
</tr>
<tr>
<td>Within BS(_1) progenies</td>
<td>rs(k−1)</td>
<td>M(_{12})</td>
<td>(\hat{\sigma}^2_{we} + \hat{\sigma}^2_{wg})</td>
</tr>
<tr>
<td>Within BS(_2) progenies</td>
<td>rt(k−1)</td>
<td>M(_{13})</td>
<td>(\hat{\sigma}^2_{we} + \hat{\sigma}^2_{wg})</td>
</tr>
<tr>
<td>Within P(_1), P(_2), F(_1)</td>
<td>rh(k−1)</td>
<td>M(_{14})</td>
<td>(\hat{\sigma}^2_{we})</td>
</tr>
</tbody>
</table>

Homogeneous errors were assumed among the F\(_3\), BS\(_1\), and BS\(_2\) progeny means, but the error mean squares can be partitioned in like manner as entries. A complete listing of genetic and error components of variance for each source of variation in Table 4.35 is as follows:

\[ \hat{\sigma}^2_{F3} = (\hat{\gamma}_2)A + (\hat{\gamma}_8)D + E_2 \] and \(\hat{\sigma}^2_{F3} = (\hat{\gamma}_4)A + (\hat{\gamma}_8)D + E_1\)
\[ \hat{\sigma}^2_{BS1} + \hat{\sigma}^2_{BS2} = (\hat{\gamma}_2)A + (\hat{\gamma}_8)D + 2E_2 \]
\[ \hat{\sigma}^2_{BS1} = (\hat{\gamma}_4)A + (\hat{\gamma}_8)D + E_1 \]
\[ \hat{\sigma}^2_{BS2} = (\hat{\gamma}_4)A + (\hat{\gamma}_8)D + E_1 \]
\[ \hat{\sigma}^2 / r = E_2 \]
\[ M_{14} = E_1 \]

The estimates of A, D, E\(_1\), and E\(_2\) are not as easily determined as when the number of equations is equal to the number of unknowns. In this example we have seven equations and four unknowns for which we desire to obtain estimates. The best procedure is to use the least-squares analysis or, more appropriately, a weighted least-squares analysis.
For maize, and particularly self-pollinated species, additional generations that increase the number of equations available for estimating components of variance can be obtained by additional generations of inbreeding, each of which provides estimates of variance among progenies and mean variance of progenies. If, for example, the F₃ and F₄ generations are included, we will have six equations and four parameters that we wish to estimate. A summary of expectations of variance among progenies and mean variance of progenies is shown in Table 4.36. For comparison, the expressions used by Mather (1949) and Mather and Jinks (1971) are shown with the equivalent expressions for $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$. Also, the equivalency of inbreeding generations is shown for the F and S designations commonly used with inbreeding. Because the F₂ generation is the reference population, the S designation is included to show the relations commonly used for broad genetic base populations. It is seen that the variance among F₂ plants is equivalent to the variance among S₀ plants. Albeit the F₂ population was obtained by selfing the F₁ hybrid, this selfing was merely to develop the F₂ population with an expected gene frequency of 0.5. Also, the variance among F₃ progenies is equivalent to variation among S₁ progenies. Genetic relations, as expected, are the same, but the differences in terminology are often confusing. The F₂ is our reference population for the special case of $p = q = 0.5$; and although it is obtained by inbreeding (selfing), its genotypic structure is equivalent to a random mating non-inbred population with the same gene frequency.

Estimates of genetic and environmental parameters estimated from F₂ populations are applicable only for the populations developed from the specific pair of inbred lines. Estimates are expected to be different if other pairs of inbred lines are used.

Estimates obtained for F₂ populations indicate relative types of variation present and future prospects of selection. But estimates usually are for only short-term objectives in comparison with populations that have a broad genetic base. Composition of components of variance in terms of additive genetic and dominance components of variance of F₂ and broad genetic base populations are equivalent if one assumes that the gene frequency of segregating loci in broad genetic base populations is 0.5. But this is not often the case in genetically broad-based populations.

Linkage effects on estimates of additive genetic and dominance variances obtained from populations developed from inbred lines are the same as those discussed for design III. As explained in previous sections, dominance variance will be biased upward regardless of linkage phase, and additive genetic variance will be biased downward for repulsion and upward for coupling phase linkages. As for design III, equilibrium of linkages can be approached by random mating of the F₂ population for four or more generations (Hanson, 1959). The F₂ and backcross populations are commonly used as base populations in applied maize breeding programs by either crossing two elite inbred lines or crossing an elite line to another line that has a specific trait to incorporate in the elite line. For F₂ populations, linkage groups may be highly desirable, whereas in backcross populations the breaking
Table 4.36  Summary of the means, variance among progenies, and mean progeny variance with continuous inbreeding and the comparison of the relations using Mather's (1949) notation

<table>
<thead>
<tr>
<th>Generation</th>
<th>Mean</th>
<th>Variance among progenies&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Mean progeny variance&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mather General</td>
<td>Mather General</td>
</tr>
<tr>
<td>$F_1$</td>
<td></td>
<td>$\hat{h} d$</td>
<td>$0 \hat{\sigma}^2 + \hat{\sigma}^2_D$</td>
</tr>
<tr>
<td>$F_2$</td>
<td>$S_0$</td>
<td>$(\frac{1}{2})\hat{h} (\frac{1}{2})d$</td>
<td>$(\frac{1}{2}D + (\frac{1}{4})H \hat{\sigma}^2_A + \hat{\sigma}^2_D$</td>
</tr>
<tr>
<td>$F_3$</td>
<td>$S_1$</td>
<td>$(\frac{1}{4})\hat{h} (\frac{1}{4})d$</td>
<td>$(\frac{1}{2}D + (\frac{1}{4})H \hat{\sigma}^2_A + (\frac{1}{2})\hat{\sigma}^2_D$</td>
</tr>
<tr>
<td>$F_4$</td>
<td>$S_2$</td>
<td>$(\frac{1}{8})\hat{h} (\frac{1}{8})d$</td>
<td>$(\frac{1}{4})D + (\frac{1}{16})H \hat{\sigma}^2_A + (\frac{3}{4})\hat{\sigma}^2_D$</td>
</tr>
<tr>
<td>$F_5$</td>
<td>$S_3$</td>
<td>$(\frac{1}{16})\hat{h} (\frac{1}{16})d$</td>
<td>$(\frac{1}{8})D + (\frac{1}{32})H \hat{\sigma}^2_A + (\frac{3}{8})\hat{\sigma}^2_D$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$F_{10}$</td>
<td>$S_8$</td>
<td>$(\frac{1}{512})\hat{h} (\frac{1}{512})d$</td>
<td>$\sim D \hat{\sigma}^2_A \sim 2\hat{\sigma}^2_A$</td>
</tr>
</tbody>
</table>

<sup>a</sup>See Table 3.2 for situations in which estimates of $\hat{\sigma}^2_D$ are valid
of tight linkages may be highly desirable and may be difficult, particularly if they have an undesirable pleiotropic effect.

Estimates of components of variance for F2 populations are important for

(1) obtaining some preliminary information on the inheritance of a trait and
(2) predicting gain for cyclical selection.

In the first instance, the two inbred parents often represent the extremes in the expression of the trait under study and effects of linkage bias could be important in estimates of genetic variation (due to additive and dominance effects) that are needed to estimate the heritability of a trait. In the second instance, long-term cyclical selection experiments usually are not conducted in F2 populations because of the limited range of genetic variability. A few cyclical selection programs, however, have been conducted to answer specific questions relative to selection theory, in which valid estimates of components of variance are necessary to predict future genetic gain and to show how predicted gain compares with observed gain (Robinson et al., 1949). Continued progress from selection in F2 populations may result because of genetic variability released from breakup of linkages in recombination generations of each cycle.

4.14 Epistasis

Epistasis was assumed to be absent in estimation of variances across mating designs, or the data were transformed to an additive scale before analysis. The assumption of no epistasis is no more serious than for other mating designs. It seems that epistasis is relatively minor compared to additive and dominance effects; most attempts to estimate epistatic variance, however, have been for broad genetic base populations. For F2 and backcross populations developed from inbred lines, epistatic effects seem to be more important than for broad genetic base populations. It is important, therefore, that the possible bias due to epistasis be checked by comparing the means of different generations, as given by Mather (1949). Estimates of components of variance and heritability are valid for the model of $p = q = 0.5$, but one must not extend the results to other populations. This restriction on interpretation of results is the same for broad genetic base populations, but the range of genotypes is more restrictive for populations developed from inbred lines than for broad genetic base populations such as open-pollinated, synthetic, and composite varieties.

It seems logical, however, that whatever the magnitude of epistatic effects, they must be present in the functioning of a genotype. Epistatic effects have been demonstrated in the expression of traits involving two loci, and thus it does not seem reasonable that epistatic effects are not operative in the expression of a complex quantitative trait such as yield. Although epistatic effects most certainly are operative, the important aspect is what proportion of the total genetic variance can be attributed to epistatic variance. If the proportion of epistatic variance to total genetic
Hereditary Variance: Mating Designs

variance is relatively small, the bias in our estimates caused by assuming no epistasis will not seriously hinder our selection progress. Buckler et al. (2009) and Holland (2009) detected only up to 2% (e.g., flowering time, a trait easily amenable to selection with simple and inexpensive methods such as stratified mass selection) of the phenotypic variation explained by marker combinations for epistatic interactions on thousands of lines derived from elite by elite combinations. Moreover, epistatic effects for yield were not detected with experiments authors emphasized that included nearly 1 million plants. They concluded that the low epistasis detected is surprising due to the presence of interactive molecular pathways and epistatic effects in other organisms. On the other hand, Dudley and Johnson (2009) suggested adding epistasis in predicting models as it significantly increased predictive power in 500 S2 lines. In this study not only lines but also testcrosses were evaluated. Of the total epistatic variance, only additive types of epistasis are fixable or usable by the breeder. Other attempts, mostly molecular approaches, to estimate epistasis are under progress and not very encouraging (Holland, 2009). However, epistasis is important from a plant breeding perspective even though reports of epistatic interactions from molecular marker data are limited (Dudley and Johnson, 2009). Choice of germplasm and accurate sampling determine the success of each project independent from the number of individuals observed.

Epistasis, as defined for quantitatively inherited traits, is purely a statistical description and does not define a physiological function or expression. Epistatic variance is an orthogonal partitioning of the interaction among loci (interallelic interaction); whereas the variability of the three phases within a locus (e.g., AA, Aa, and aa) is partitioned into 1 df for sums of squares due to regression ($\hat{\sigma}^2_A$) and 1 df for the deviations from regression (dominance, $\hat{\sigma}^2_D$, or interaction within a locus). Hence epistasis, in a statistical sense, is the non-additivity of effects among loci as contrasted with dominance effects, which are due to non-additivity within a locus.

Cockerham (1954) developed the relations for partitioning total epistatic variance into the three types of epistasis: additive $\times$ additive, additive $\times$ dominance, and dominance $\times$ dominance for a two-locus model. Higher order epistasis follows by including additional loci. Cockerham (1956a) illustrated different genetic models for the expression of different types of epistasis for a two-locus situation. He also suggested how epistasis may be estimated by use of different mating designs and different inbreeding levels of parents included in these designs. Combinations of mating designs and levels of inbreeding provide additional equations and differences in coefficients of genetic components of variance for covariances of relatives.

Two attempts to estimate epistatic variance in maize populations were reported by Eberhart et al. (1966) and Silva and Hallauer (1975). In both instances, progenies developed by designs I and II were evaluated, where the parents were non-inbred ($F = 0$) for design I matings and inbred ($F = 1$) for design II matings. The inbred parents were unselected inbred lines developed from the same populations from which the non-inbred parents were sampled. Hence two samplings of $S_0$ plants were made from the population: (1) $S_0$ plants used as non-inbred parents of design I matings and (2) $S_0$ plants that were the progenitors of the unselected inbred lines. Because expectations of covariances of relatives are different for the two levels of inbreeding, additional equations are available for estimation. Tables 4.37 and 4.38 give examples of coefficients of components of variance for an experiment.
Table 4.37 Expected components of genetic variance and coefficients for design I and II mating designs

<table>
<thead>
<tr>
<th>Components of variance</th>
<th>( \hat{\sigma}^2_A )</th>
<th>( \hat{\sigma}^2_D )</th>
<th>( \hat{\sigma}^2_{AA} )</th>
<th>( \hat{\sigma}^2_{AD} )</th>
<th>( \hat{\sigma}^2_{DD} )</th>
<th>( \hat{\sigma}^{2}_e )</th>
<th>( \hat{\sigma}^{2}_w )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design I (F = 0)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\sigma}^2_m )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{16} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\sigma}^2_{fm} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{3}{16} )</td>
<td>( \frac{1}{8} )</td>
<td>( \frac{1}{16} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\sigma}^2 )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{3}{4} )</td>
<td>( \frac{3}{4} )</td>
<td>( \frac{7}{8} )</td>
<td>( \frac{15}{16} )</td>
<td>( 1 )</td>
<td></td>
</tr>
<tr>
<td>( \hat{\sigma}^2_w )</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{2} )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
</tr>
</tbody>
</table>

| **Design II (F = 1)** |                |                |                |                |                |                |                |
| \( \hat{\sigma}^2_m \) | \( \frac{1}{2} \) | \( \frac{1}{4} \) |                |                |                |                |                |
| \( \hat{\sigma}^2_f \) | \( \frac{1}{2} \) | \( \frac{1}{4} \) |                |                |                |                |                |
| \( \hat{\sigma}^2_{fm} \) | \( 1 \)       | \( \frac{1}{2} \) | \( 1 \)       | \( 1 \)       | \( 1 \)       | \( 1 \)       | \( 1 \)       |
| \( \hat{\sigma}^2 \)   | \( \frac{1}{2} \) | \( \frac{1}{4} \) | \( \frac{1}{2} \) | \( 1 \)       | \( 1 \)       | \( 1 \)       | \( 1 \)       |
| \( \hat{\sigma}^2_w \) | \( \frac{1}{2} \) | \( \frac{1}{4} \) | \( \frac{1}{2} \) | \( 1 \)       | \( 1 \)       | \( 1 \)       | \( 1 \)       |

that includes design I and II progenies. Coefficients of components of variance in Table 4.37 were obtained from expected mean squares. Nine equations are available, and a genetic model that includes digenic epistatic components could be fitted by least-squares analysis.

Table 4.38 Expectation of mean squares in terms of genetic and environmental variance components pooled over sets when within-plot variances are available for design I and II mating designs

<table>
<thead>
<tr>
<th>Source</th>
<th>( \hat{\sigma}^2_A )</th>
<th>( \hat{\sigma}^2_D )</th>
<th>( \hat{\sigma}^2_{AA} )</th>
<th>( \hat{\sigma}^2_{AD} )</th>
<th>( \hat{\sigma}^2_{DD} )</th>
<th>( \hat{\sigma}^2_e )</th>
<th>( \hat{\sigma}^2_w )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design I</strong> ( r = 2, f = 6 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (M)/sets (S)</td>
<td>3.554</td>
<td>0.582</td>
<td>1.207</td>
<td>0.345</td>
<td>0.227</td>
<td>0.109</td>
<td>1</td>
</tr>
<tr>
<td>Females/M/S</td>
<td>0.554</td>
<td>0.582</td>
<td>0.457</td>
<td>0.345</td>
<td>0.227</td>
<td>0.109</td>
<td>1</td>
</tr>
<tr>
<td>Error</td>
<td>0.054</td>
<td>0.082</td>
<td>0.082</td>
<td>0.095</td>
<td>0.102</td>
<td>0.109</td>
<td>1</td>
</tr>
<tr>
<td>Within plot</td>
<td>0.500</td>
<td>0.750</td>
<td>0.750</td>
<td>0.875</td>
<td>0.938</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td><strong>Design II</strong> ( r = 2, m = 4, f = 4 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (M)/sets (S)</td>
<td>4.000</td>
<td>2.000</td>
<td>3.000</td>
<td>2.000</td>
<td>2.000</td>
<td>0.110</td>
<td>1</td>
</tr>
<tr>
<td>Females (F)/S</td>
<td>4.000</td>
<td>2.000</td>
<td>3.000</td>
<td>2.000</td>
<td>2.000</td>
<td>0.110</td>
<td>1</td>
</tr>
<tr>
<td>(M × F)/S</td>
<td>0.000</td>
<td>2.000</td>
<td>1.000</td>
<td>2.000</td>
<td>2.000</td>
<td>0.110</td>
<td>1</td>
</tr>
<tr>
<td>Error</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.110</td>
<td>1</td>
</tr>
<tr>
<td>Within plot ( c )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0</td>
</tr>
</tbody>
</table>

\( c \) Harmonic means of the number of plants per plot \( k \) were \( k = 9.2 \) for design I and \( k = 9.1 \) for design II.
Although we have sufficient equations for estimation of epistatic variance, general results have not been encouraging. Eberhart et al. (1966) and Silva and Hallauer (1975) were unable to obtain realistic estimates of digenic epistatic components. Silva and Hallauer, for example, found that $\hat{\sigma}^2_A$ accounted for 93.2% of total genetic variance for yield and that inclusion of $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$ in the model accounted for 99% of total genetic variation, with no improvement in the fit when $\hat{\sigma}^2_{AA}$ was included. Because the mean squares have unequal variances, it was considered that weighted least squares or maximum likelihood methods of estimation would be an improvement. But when the models included more than one digenic epistatic component they usually were negative and unrealistic, with much greater standard errors. A model that included as many terms as permitted by the number of independent equations made the X-matrix nearly singular.

Dudley and Johnson (2009) added epistasis to a partial least-squares model and increased predicting power in grain yield and quality. They suggested evaluating all possible marker interactions rather than only those between significant markers as quantitative traits probably are part of a complex gene network involving segments across all chromosomes. The authors caution most breeding populations carry unbroken linkage blocks that may contain blocks of genes for which interactions will average out making epistatic models not useful in genetically narrow-based populations. Large sample sizes and extensive testing across environments were encouraged to increase predicting power. Repeating the experiments over environments also provides additional equations that can be used to estimate interaction of genetic effects with environments.

It was previously shown that a series of $F$-tests can be made for triallel and quadrallel analyses of variance to test hypotheses that include epistasis. If the $F$-tests indicate significant variation for different sources of variation, one can proceed to estimate genetic components of variance, including digenic epistasis. For instance, in the triallel analysis there are nine covariances of relatives that have different genetic expectations. Solutions of the set of simultaneous linear equations of the mean squares (or components of variance) will then permit the fitting of genetic models that include trigenic additive epistatic variance. Wright et al. (1971) used triallel and diallel analyses to estimate epistasis by use of 60 unselected inbred lines developed from a selected strain of Krug Yellow Dent, which is an open-pollinated variety. From pooled analyses of the triallel and diallel, nine linear equations were available to fit genetic models that included trigenic additive epistatic variance. Unweighted least squares and maximum likelihood estimation procedures were compared for estimation of genetic components of variance. Regardless of genetic model and estimation procedure used, $\hat{\sigma}^2_A$ accounted for the largest proportion of total genetic variance. Fitting the error and a six-parameter genetic model ($\hat{\sigma}^2_A, \hat{\sigma}^2_D, \hat{\sigma}^2_{AA}, \hat{\sigma}^2_{DD}, \hat{\sigma}^2_{AAA}$) showed that it was not possible to obtain realistic estimates of epistatic components of variance, although significant epistatic effects were detected in triallel analyses of variance. Negative estimates of epistatic variance components were frequent in the unweighted analysis, but standard errors of estimates indicated the estimates were
within the range of zero. Negative estimates of epistatic components also were frequent by the maximum likelihood procedure, but they were smaller and usually had smaller standard errors. Only 0.27% and 0.24% of total genetic variation of yield for unweighted and weighted estimation procedures, respectively, were attributed to the four epistatic components of variance; whereas fitting the error and $\hat{\sigma}^2_A$ accounted for 99 and 97% of total variation for yield.

Chi et al. (1969) used a complex mating design in the Reid Yellow Dent open-pollinated variety that included 66 covariances within and among 11 branches of a family, which represented five broad categories of relatives: full-sibs, half-sibs, cousins, uncle-nephew, and no relationship. Seven genetic models were fitted to observe mean squares that included up to trigenic additive epistasis. They used unweighted least-squares analyses and found no evidence of epistasis in the Reid Yellow Dent variety. Estimates from a model that included $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ indicated that the major part of the total genetic variance could be attributed to these two parameters. Chi et al. (1969) also have indicated that one of the main problems involves coefficients of digenic and trigenic epistatic components of variance because they are highly correlated with those of additive and dominance variance components. The correlation occurs because the coefficients of epistatic components of variance are generated by either squaring or multiplying the coefficients of additive and dominance variance components. This is evident from the general relations for covariances of relatives shown in Chapter 2. Hence this inherent property of the model for covariances of relatives reduces the sensitivity of the model for detecting epistasis. For the model that included six components of genetic variance used by Chi et al. (1969), the correlation matrix of coefficients of genetic components of variance is shown in Table 4.39.

The correlations of coefficients among epistatic components of variance and of $\hat{\sigma}_D^2$ with the epistatic components are very high in all instances, e.g., $\hat{\sigma}_{AA}^2$ has a correlation of 0.92 with $\hat{\sigma}_A^2$. Consequently, high correlations of the coefficients will give greater standard errors of estimates of genetic components of variance.

Although large studies have been conducted, including a sufficient number of mean squares for the fitting of genetic models that included epistasis, they all failed to obtain realistic estimates of epistatic components of variance. In all instances, except for Dudley and Johnson (2009) at increased significant levels, the reduced

### Table 4.39 Correlation matrix of coefficients of genetic components of variance

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\sigma}_A^2$</th>
<th>$\hat{\sigma}_D^2$</th>
<th>$\hat{\sigma}_{AA}^2$</th>
<th>$\hat{\sigma}_{AD}^2$</th>
<th>$\hat{\sigma}_{DD}^2$</th>
<th>$\hat{\sigma}_{AAA}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_A^2$</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_D^2$</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_{AA}^2$</td>
<td>0.92</td>
<td>0.93</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_{AD}^2$</td>
<td>0.71</td>
<td>0.98</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_{DD}^2$</td>
<td>0.67</td>
<td>0.95</td>
<td>0.89</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_{AAA}^2$</td>
<td>0.81</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
<td>0.97</td>
<td>1.00</td>
</tr>
</tbody>
</table>
models that included error, \( \hat{\sigma}^2_A \), and \( \hat{\sigma}^2_D \) accounted for most of the total variation. It seems, in general, that the relative proportion of total genetic variance due to epistatic variance is quite small.

Several assumptions were used in obtaining estimates of components of genetic variance: (1) randomness of parents, (2) diploid inheritance, (3) no linkage effects or linkage equilibrium, (4) no maternal effects, and (5) additive environmental and genotypic effects. The studies of Eberhart et al. (1966), Wright et al. (1971), and Silva and Hallauer (1975) used unselected inbred lines developed from populations. In all instances selection was minimized, but it is also obvious that it was impossible to maintain the progeny of each original S_0 plant during the inbreeding process. Eberhart et al. (1966) and Wright et al. (1971) had a relatively small sample of inbred lines developed from the original S_0 plants sampled. Silva and Hallauer (1975) had a sample of 160 inbred lines, but the estimation of epistatic variance was no more successful than those of Eberhart et al. (1966) and Wright et al. (1971). Increasing the number of progenies to 500 S_2 lines and testcrosses from Illinois High × Low Oil and Illinois High × Low Protein seems to make a difference even though probability levels were modified. Wolf et al. (2000) evaluated S_1 progenies and testcross means for the DIII mating design that included 100 F_2 male plants of the (B73 x M_O17) cross. Although 10 mean squares and mean products were available for estimation of genetic components of variance through digenic epistasis, Wolf et al. (2000) reported estimates of epistatic variances were relatively less important than estimates additive genetic and dominance variances for the (B73 x M_O17) F_2 population. The effects of linkage could be a disturbing element, and Cockerham (1956b) and Schnell (1963) have shown that linkage effects increase the coefficients of epistatic terms of covariances of relatives. In all studies except Wright et al. (1971), Buckler et al. (2009), and Dudley and Johnson (2009), the parents were derived from populations that should approach linkage equilibrium. Complete linkage of all factors controlling a trait on each chromosome does not seem probable in maize. Linkage bias on epistatic components would not be of much importance unless the average recombination frequency was 0.1 or less. Hence effects of linkage should not have been an important source of bias in all studies unless interactions average out (Dudley and Johnson, 2009). Assumptions of diploid inheritance, no maternal effects, and additivity of genetic and environmental effects have not been found to be invalid in maize. The small coefficients of epistatic terms relative to the coefficients of additive and dominance terms are an important factor in estimation of epistatic components of variance. The smallness of coefficients also contributes to greater standard errors of epistatic estimates.

Qualitative evidence of epistatic effects, however, has been reported from comparison of means of different types of hybrids. In most instances the hybrids were produced from inbred lines and desirable epistatic combinations would be fixed in those lines. The use of the same set of inbred lines to produce a balanced set of single, three-way, and double-cross hybrids would indicate effects of recombination for epistatic effects. Recombination would be present in the one single cross used to produce the three-way crosses and in both single crosses for production of the double-cross hybrids. Comparisons of different permutations of crosses for the same set of inbred lines give qualitative evidence of the presence of epistatic effects.
for specific combinations of inbred lines. This type of evidence for epistatic effects cannot be quantified to determine its relative importance to other types of genetic variance, but it demonstrates that epistatic effects are important for specific combinations of inbred lines. Holland (2001) has presented a comprehensive review for the formulae and methods used to detect and estimate epistasis, possible effects on inbreeding depression, heterosis, and response to selection, and the implications of epistasis in plant breeding. Holland (2009) has also presented the challenges for detecting epistasis with molecular markers and their interactions.

References


Hanson, W. D. 1959. The breakup of initial linkage blocks under selected mating systems. *Genetics* 44:857–68.


References

Chapter 5
Hereditary Variance: Experimental Estimates

The mating designs described in Chapter 4 have been used extensively in maize to determine relative proportions of total variation that are governed by genetic and environmental forces and to characterize genetic variation due to additive and non-additive effects. Maize is amenable to study by the different mating designs because of the ease in obtaining sufficient quantities of seed for testing by cross- and self-fertilization. Because maize is a naturally cross-fertilized crop species, variability within maize populations was obvious to researchers. Until the late 1940s maize breeders and researchers emphasized the development of procedures for increasing effectiveness and efficiency of inbred line and hybrid development based on the principles given by East (1908), Shull (1908, 1909), and Jones (1918). Because of the variability within populations maize breeders, unlike animal breeders, did not concern themselves with attempting to characterize the types of genetic variability present and how this could influence effectiveness of selection of lines and their expression in hybrids until the publications of Comstock and Robinson (1948) and Mather (1949). Similarly, improvement of the populations was generally ignored. The papers by Jenkins (1940), Hull (1945), and Comstock et al. (1949) integrated possible effects of types of gene action on efficiency of selection and stimulated interest in maize populations and their improvement by breeders.

Since the 1940s, researchers have been very active in estimation of genetic and environmental components of variance for different types of maize populations. Additionally, they have attempted to determine the relative proportions of total genetic variance that are attributable to additive and non-additive effects, both intra- and inter allelic. Voluminous literature has developed, presenting the empirical results of these studies and how they compare with theory and different selection procedures. Several different population types have been sampled because of the interest in possible differences of genetic variability among populations. Because of commitment to hybrid use the ratio of dominance variance to additive genetic variance has received considerable attention and has been estimated in several studies. The F2 populations developed from a cross of two inbred lines have been studied in some instances to estimate specific parameters; e.g., gene frequency expected to be 0.5 for all segregating loci permits estimation of the average level of dominance of genes affecting the trait under study. Also, F2 populations and their advanced
generations have been studied to estimate effects of linkages on estimates of additive and dominance variances.

5.1 Experimental Results

Table 5.1 summarizes the estimates of additive variance ($\hat{\sigma}_A^2$) and dominance variance ($\hat{\sigma}_D^2$) available for 19 different traits of maize. Most estimates were obtained by use of mating designs I, II, and III, but few estimates for F2 populations were obtained (Mather, 1949). No estimates from diallel analyses were included. The last column of Table 5.1 (No.) shows the number of estimates included for each trait as the precision of estimates is determined by the number of estimates included in the average. Estimates of components of variance shown in Table 5.1 are averages for each trait.

The greatest number of estimates was reported for yield, and averages for each parameter show that the ratio of dominance to additive variance was quite large for yield when compared to other traits. Therefore, dominance variance seems important in the expression of yield. If estimates of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ for each study are totaled and divided by the total number of estimates, we obtain 0.9377 (Table 5.1). If, however, we divide the average estimate of the dominance variance by the average estimate of the additive variance, we obtain 0.6113, shown in parentheses in Table 5.1. The difference between the two estimates of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ arises because of the frequent occurrence of negative estimates of $\hat{\sigma}_D^2$. When the estimate of $\hat{\sigma}_D^2$ was reported as a negative value, it was consequently not possible to determine the ratio of $\hat{\sigma}_D^2/\hat{\sigma}_A^2$. An unbiased average of $\hat{\sigma}_D^2$ was obtained by including negative estimates in the algebraic summation of estimates of $\hat{\sigma}_D^2$ from the reported studies. If estimates of $\hat{\sigma}_D^2$ are in reality either very small positive values or zero, negative experimental estimates are not unexpected. Hence the average of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ would be biased upward because the ratios in individual studies that had negative estimates of $\hat{\sigma}_D^2$ were not available; if the true estimates were very small positive values, the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ would be small. The true value of a parameter for a particular trait for a particular population will be approached by repeated experimentation; thus the best estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ are those averaged across all the reported studies, provided we make the broad generalization that maize is our population. It seems, therefore, that the best estimate of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ is the one determined from the average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. Although the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ calculated from the averages of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ is lower than the average of ratios, considerable dominance variance is expressed for yield. Assuming no epistasis and linkage effects, $\hat{\sigma}_A^2$ on the average accounted for 61.2% and $\hat{\sigma}_D^2$ accounted for 38.8% of the total genetic variation for yield.

The standard errors of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ also are included in Table 5.1. Relative to $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$, standard errors for yield were much larger for $\hat{\sigma}_D^2$. Many of the estimates for $\hat{\sigma}_D^2$ were obtained from studies that used the design I mating scheme, and it is known that the standard errors for these estimates are large because of the complex
### Table 5.1

Summary for 19 maize traits for the average estimates of additive (\(\hat{\sigma}^2_A\)) and dominance (\(\hat{\sigma}^2_D\)) components of variance and their standard errors (SE), ratio of dominance to additive genetic variance (\(\hat{\sigma}^2_D/\hat{\sigma}^2_A\)), and heritability (\(\hat{h}^2\)) on a plot basis.

<table>
<thead>
<tr>
<th>Trait</th>
<th>(\hat{\sigma}^2_A)</th>
<th>SE((\hat{\sigma}^2_A))</th>
<th>(\hat{\sigma}^2_D)</th>
<th>SE((\hat{\sigma}^2_D))</th>
<th>(\hat{\sigma}^2_D/\hat{\sigma}^2_A)</th>
<th>(\hat{h}^2) (%)</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield (g)(^a)</td>
<td>469.1</td>
<td>174.3</td>
<td>286.8</td>
<td>210.1</td>
<td>0.9377</td>
<td>(0.6113)(^b)</td>
<td>18.7 99</td>
</tr>
<tr>
<td>Plant height (cm)</td>
<td>212.9</td>
<td>51.6</td>
<td>36.2</td>
<td>46.5</td>
<td>0.5338</td>
<td>(0.1700)</td>
<td>56.9 45</td>
</tr>
<tr>
<td>Ear height (cm)</td>
<td>152.7</td>
<td>35.5</td>
<td>11.1</td>
<td>36.5</td>
<td>0.3743</td>
<td>(0.2324)</td>
<td>66.2 52</td>
</tr>
<tr>
<td>Number of ears ((\times 10^3))</td>
<td>45.9</td>
<td>13.2</td>
<td>11.8</td>
<td>9.9</td>
<td>0.4366</td>
<td>(0.2875)</td>
<td>39.0 39</td>
</tr>
<tr>
<td>Ear length (cm) ((\times 10^2))</td>
<td>152.4</td>
<td>37.8</td>
<td>50.4</td>
<td>47.3</td>
<td>0.3746</td>
<td>(0.2480)</td>
<td>38.1 36</td>
</tr>
<tr>
<td>Ear diameter (cm) ((\times 10^2))</td>
<td>4.6</td>
<td>1.1</td>
<td>0.9</td>
<td>1.1</td>
<td>0.3269</td>
<td>(0.2391)</td>
<td>36.1 35</td>
</tr>
<tr>
<td>Kernel-row number ((\times 10^2))</td>
<td>189.0</td>
<td>45.5</td>
<td>14.5</td>
<td>77.8</td>
<td>0.1774</td>
<td>(0.2407)</td>
<td>57.0 18</td>
</tr>
<tr>
<td>Kernel weight (g)</td>
<td>34.9</td>
<td>8.5</td>
<td>9.5</td>
<td>9.4</td>
<td>0.5544</td>
<td>(0.2435)</td>
<td>41.8 11</td>
</tr>
<tr>
<td>Days to flower</td>
<td>4.0</td>
<td>0.9</td>
<td>-0.1</td>
<td>0.9</td>
<td>0.6598</td>
<td>(---)</td>
<td>57.9 48</td>
</tr>
<tr>
<td>Grain moisture (%)</td>
<td>7.2</td>
<td>1.7</td>
<td>0.5</td>
<td>2.5</td>
<td>0.4801</td>
<td>(0.2361)</td>
<td>62.0 4</td>
</tr>
<tr>
<td>Oil (%) ((\times 10^2))</td>
<td>82.2</td>
<td>15.6</td>
<td>8.7</td>
<td>8.8</td>
<td>0.1808</td>
<td>(0.1897)</td>
<td>76.7 4</td>
</tr>
<tr>
<td>Lodging (%) ((\times 10^3))</td>
<td>126.1</td>
<td>33.6</td>
<td>-30.2</td>
<td>24.2</td>
<td>0.0265</td>
<td>(---)</td>
<td>71.9 5</td>
</tr>
<tr>
<td>Number of tillers ((\times 10^2))</td>
<td>26.9</td>
<td>6.0</td>
<td>-1.6</td>
<td>---</td>
<td>0.1850</td>
<td>(---)</td>
<td>---</td>
</tr>
<tr>
<td>Kernel depth ((\times 10^3))</td>
<td>18.7</td>
<td>4.2</td>
<td>5.0</td>
<td>3.6</td>
<td>0.5114</td>
<td>(0.2673)</td>
<td>29.2 7</td>
</tr>
<tr>
<td>Cob diameter ((\times 10^2))</td>
<td>16.6</td>
<td>2.8</td>
<td>3.4</td>
<td>3.0</td>
<td>0.2131</td>
<td>(0.2048)</td>
<td>37.0 6</td>
</tr>
<tr>
<td>Husk extension ((\times 10^2))</td>
<td>54.8</td>
<td>10.4</td>
<td>25.2</td>
<td>14.7</td>
<td>0.4598</td>
<td>(0.2452)</td>
<td>49.5 3</td>
</tr>
<tr>
<td>Husk score ((\times 10^2))</td>
<td>65.2</td>
<td>1.0</td>
<td>20.4</td>
<td>12.9</td>
<td>0.3128</td>
<td>(0.1700)</td>
<td>35.9 3</td>
</tr>
<tr>
<td>Flag leaf number ((\times 10^2))</td>
<td>67.8</td>
<td>---</td>
<td>18.0</td>
<td>---</td>
<td>0.2654</td>
<td>(---)</td>
<td>---</td>
</tr>
<tr>
<td>Flag leaf length ((\times 10^2))</td>
<td>154.0</td>
<td>---</td>
<td>58.6</td>
<td>---</td>
<td>0.3805</td>
<td>(---)</td>
<td>---</td>
</tr>
</tbody>
</table>

\(^a\)Units on per plant basis  
\(^b\)Ratio is for average estimates of \(\hat{\sigma}^2_A\) and \(\hat{\sigma}^2_D\)
individual experiments the standard error of \( \hat{\sigma}_D^2 \) was often greater than the estimate of \( \hat{\sigma}_D^2 \). However, this value was reduced in mating designs that allowed a direct estimate.

We want to emphasize that the ratio \( \hat{\sigma}_D^2/\hat{\sigma}_A^2 \) was considerably lower for other traits than for yield. In most instances the ratio when calculated from the average of estimates \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) was even lower than the average of the reported ratios \( \hat{\sigma}_D^2/\hat{\sigma}_A^2 \). It seems, therefore, that the greatest proportion of total genetic variance can be attributed to additive effects for most traits. A relatively large average ratio was obtained for days to silk; however, the average estimate for \( \hat{\sigma}_D^2 \) is a small negative value. Similar to yield, frequent negative estimates of \( \hat{\sigma}_D^2 \) were obtained, and the ratio \( \hat{\sigma}_D^2/\hat{\sigma}_A^2 \) was not estimable. Hence the average of the ratios is biased upward because only the ratios that had positive estimates of \( \hat{\sigma}_D^2 \) were included. Similarly, the best estimate of the ratio seems to be the one from the unbiased average of estimates of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \), which for days to silk must be near zero. The average ratio \( \hat{\sigma}_D^2/\hat{\sigma}_A^2 \) for kernel weight and kernel depth indicated similar estimates of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \). In both instances, ratios were considerably lower when unbiased averages of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) were compared. Average estimates of \( \hat{\sigma}_A^2 \) were about four times greater than average estimates of \( \hat{\sigma}_D^2 \).

Estimates of heritability were calculated on a per plot basis, where the necessary components of variance were given in the reported studies. As Hanson (1963) has emphasized, the unit used for reporting heritability estimates is very important in plant research. Because several of the studies were conducted in only one environment and did not include genotype–environment components of variance, estimates of heritability were calculated as

\[
\hat{h}^2 = \hat{\sigma}_A^2/(\hat{\sigma}_D^2 + \hat{\sigma}_A^2)
\]

\( \hat{\sigma}^2 \) is the plot experimental error.

If the estimates of the components of variance for genetic variability, their interactions with environments, and experimental error had been available in all instances, heritability on a progeny mean basis would be more meaningful. Estimates of heritability for the 16 traits that reported positive values in Table 5.1 were distributed according to the following ranges:

<table>
<thead>
<tr>
<th>Heritability (%)</th>
<th>Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{h}^2 &gt; 70 )</td>
<td>Percent oil, number of tillers</td>
</tr>
<tr>
<td>( 50 &lt; \hat{h}^2 &lt; 70 )</td>
<td>Plant height, ear height, kernel-row number, days to flower, grain moisture</td>
</tr>
<tr>
<td>( 30 &lt; \hat{h}^2 &lt; 50 )</td>
<td>Number of ears, ear length, ear diameter, kernel weight, husk extension, husk score, cob diameter</td>
</tr>
<tr>
<td>( \hat{h}^2 &lt; 30 )</td>
<td>Yield and kernel depth</td>
</tr>
</tbody>
</table>
5.1 Experimental Results

The magnitude of average heritability estimates reflects the number of estimates reported in the literature and the complexity of the traits. Yield is the most economically important trait in maize, and its heritability is the lowest of all traits. Yield also had a relatively large proportion of the total variance accounted for by $\hat{\sigma}_D^2$. Kernel depth, number of ears, ear length, and ear diameter are components of yield, and their estimates of heritability were about twice as large as that for yield. Yield results from the total expression of the genotype from the time seed is planted until harvest. As a consequence, yield itself is the combined expression of genotype and environment throughout the duration of the growing season. Yield components, however, are determined during certain stages of the ontogeny of the genotype, so their expression depends on just a portion of the growing season. Number of ears per plant, for example, is determined by the combination of genetic and environmental forces from 6 weeks before flowering to flowering time. If environmental conditions and genetic composition of the genotype are favorable for more than one ear per plant before and at flowering time, the plants will have more than one ear. If the combination of conditions is unfavorable for more than one ear before and at flowering, more than one ear will not result even though optimum conditions may occur for ear development and grain fill after flowering. Kernel depth, ear length, and ear diameter are yield components but influenced to a large extent by environmental conditions after flowering.

Kernel-row number is another yield component, but it has a relatively high average heritability estimate (57.0%). Kernel-row number also is determined in relatively early stages of plant ontogeny and is less affected by environmental conditions at the time of flowering and from flowering to maturity; i.e., environmental forces can affect depth of kernels, kernel weight, and size of ear development after flowering, but the number of kernel rows is not altered. Days to flowering and grain moisture at harvest are measures of maturity of maize, and they both have relatively large average heritability estimates. The relatively large heritability values for maturity are in concert with experiences of applied breeding programs because it has been relatively easy to change the maturity of maize, as evidenced by range of adaptation of maize for latitude and elevation. At the molecular level, however, flowering time seems to be a complex trait characterized by small additive QTLs with few genetic interactions (Buckler et al., 2009).

Because the heritability estimates for the yield components (kernel depth, ear length, number of kernel rows, seed size, number of kernels per row, prolificacy) were generally greater than grain yield itself, there were some suggestions that emphasis on the components of yield would be more effective for increasing grain yield itself. Studies have demonstrated that selection that emphasized only one trait was generally not effective for increasing grain yield (Geadelmann and Peterson, 1978; Odhiambo and Compton, 1987; Carena et al., 1998; Hallauer et al., 2004). Because of the complexity and number of components that affect final grain yield expression, it seems selection based only on grain (e.g., in addition to other agronomic traits) should be measured. Because of the myriad of environmental effects that affect plant and ear expression throughout the growing season, adjustments in development and growth are made by the maize plant for the composite of traits that affect the final expression of grain yield.
Average estimates of the parameters shown in Table 5.1 also were calculated for five types of maize populations: (1) F2 populations developed from a cross of two inbred lines; (2) synthetics developed from recombination of elite inbred lines (usually 10 – 24 lines); (3) open-pollinated varieties; (4) variety crosses produced by crossing either open-pollinated varieties or synthetic varieties; and (5) composites developed by inter-mating material of diverse origin, such as open-pollinated and synthetic varieties, hybrids, and inbred lines. There are large differences for average estimates for the five types of maize populations (Table 5.2); and again, the best estimates for each type of population depend on the number of experimental estimates included. The last column and last row of Table 5.2 show the number of estimates available for the five types of populations and for each parameter, respectively. For example, 37 estimates of $\hat{\sigma}_A^2$ were reported for open-pollinated varieties vs. only 10 estimates for composites. The reasons for the seeming discrepancy in the number of estimates available for each parameter are that (1) standard errors were calculated for $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$, (2) estimates of $\hat{\sigma}_D^2$ were not available from the use of inbred lines, (3) negative estimates did not permit the estimation of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$, and (4) experimental error was not included in the literature to permit calculation of heritability on a plot basis.

Estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ were largest for composite and F2 populations and smallest for synthetic varieties. Standard errors of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ also were largest for composite and F2 populations. The average estimate of $\hat{\sigma}_A^2$ for composites was greater than for F2 populations, but the average estimate of $\hat{\sigma}_D^2$ was greater for F2 populations. The average ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ was greater than one for F2, variety cross, and

Table 5.2 Summary of estimates of additive dominance and genetic components of variance, ratio of dominance variance to additive variance, and heritability estimates for yield in five types of maize populations

<table>
<thead>
<tr>
<th>Type of populations</th>
<th>$\hat{\sigma}_A^2$ (SE)</th>
<th>$\hat{\sigma}_D^2$ (SE)</th>
<th>$\hat{\sigma}_D^2/\hat{\sigma}_A^2$ (SE)</th>
<th>$\hat{h}^2$</th>
<th>$\hat{\sigma}_D^2/\hat{\sigma}_D^2$ (SE)</th>
<th>No. of reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2</td>
<td>585.1 (338.5)</td>
<td>451.0 (593.0)</td>
<td>1.0022 (0.7708)</td>
<td>24.4 (1.30)</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Synthetics</td>
<td>225.9 (59.3)</td>
<td>128.6 (83.4)</td>
<td>0.8255 (0.5690)</td>
<td>22.9 (1.76)</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Open pollinated</td>
<td>503.8 (178.9)</td>
<td>245.8 (320.8)</td>
<td>0.7619 (0.4879)</td>
<td>18.9 (2.16)</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Variety crosses</td>
<td>306.2 (139.2)</td>
<td>292.2 (32.0)</td>
<td>1.3854 (0.9540)</td>
<td>13.4 (1.05)</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Composites</td>
<td>721.9 (432.0)</td>
<td>281.8 (—)</td>
<td>1.3335 (0.3902)</td>
<td>13.8 (2.56)</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>468.6 (229.6)</td>
<td>279.9 (257.3)</td>
<td>0.9377 (0.6343)</td>
<td>18.7 (1.76)</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>No. of estimates</td>
<td>99 (55)</td>
<td>82 (41)</td>
<td>72 (43)</td>
<td>82 (99)</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

* Ratio is for the average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$
composite populations and less than one for synthetic and open-pollinated variety populations. In all instances, however, the ratios estimated from unbiased averages of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) were less than one but still relatively large for F2 and variety cross populations.

For each of the five types of populations, the average \( \hat{\sigma}_A^2 \) value was greater than the average \( \hat{\sigma}_D^2 \) value, ranging from 2.56 times greater for composite populations to 1.05 times greater for variety cross populations. Additive genetic variance \( \hat{\sigma}_A^2 \) is considerably greater than dominance variance \( \hat{\sigma}_D^2 \) for synthetic, open-pollinated, and composite populations, indicating that \( \hat{\sigma}_A^2 \) is more important for yield than \( \hat{\sigma}_D^2 \) in these populations. Average standard errors were particularly large for average estimates of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) for F2 populations and also large for average estimates of \( \hat{\sigma}_D^2 \) for open-pollinated varieties. More estimates from classical and molecular studies (e.g., both types of studies together if possible) are needed as improvements in experimental techniques, experimental designs, and analysis will be important in contributing to smaller errors in experiments.

Figure 5.1 shows the distribution of additive genetic variance experimental estimates for yield across the five types of populations. Before plotting each estimate of \( \hat{\sigma}_A^2 \), the arrangement of the five types of populations was made on the basis of expected relative magnitude of estimates of \( \hat{\sigma}_A^2 \); i.e., F2 populations have the least variability and composites the greatest. Figure 5.1, however, shows that distributions of individual estimates of \( \hat{\sigma}_A^2 \) are as great for F2 populations as for open-pollinated and composite variety populations.

Averages of \( \hat{\sigma}_A^2 \) (solid lines) and of standard errors of \( \hat{\sigma}_A^2 \) (dashed lines) also are included in Fig. 5.1. Except for the synthetic variety and variety crosses, few estimates are considerably different from the majority. Study of individual experiments from which extremely large estimates of \( \hat{\sigma}_A^2 \) were reported also shows very large estimates of standard errors of \( \hat{\sigma}_A^2 \). One negative estimate \((-144)\) of \( \hat{\sigma}_A^2 \) also was reported in an open-pollinated variety.

The most surprising feature of the distributions is the difference between distributions of F2 and synthetic variety populations. The F2 populations seem to have greater variability on the average than synthetic varieties. Because synthetic varieties were developed by recombining elite inbred lines, it would seem that variability within synthetic varieties should be at least equivalent to variability of F2 populations. Synthetic varieties would be equivalent to advanced generations of multiple F2 populations. The standard errors of estimates of \( \hat{\sigma}_A^2 \) for synthetic varieties, however, were much smaller (5.7 times) than those for F2 populations. Most of the estimates for synthetic varieties also were obtained since 1960 by use of mating designs other than design I. The greater estimates of \( \hat{\sigma}_A^2 \) for F2 populations compared with synthetic varieties may be a function of gene frequency. Assuming a genetic model of partial to complete dominance, estimates of \( \hat{\sigma}_A^2 \) for F2 populations whose allele frequencies at the segregating loci would be 0.5 would be greater than \( \hat{\sigma}_A^2 \) in synthetic varieties which may have allele frequencies greater than 0.5. Synthetic varieties usually were developed by intermating elite inbred lines. Also, allele frequencies
in open-pollinated varieties may be near 0.5 or lower and, with partial to complete dominance, $\hat{\sigma}^2_A$ might be near the maximum. Estimates of $\hat{\sigma}^2_A$ for variety cross populations have definitions slightly different from those for intra-populations and may not be equivalent.

Because estimates of $\hat{\sigma}^2_A$ are gene-frequency dependent, differences among the five types of populations may be influenced as much by frequency of segregating loci as by number of segregating loci. From Fig. 5.1 it seems that sufficient variability is present in any type of population to expect significant progress from selection. The maize breeder must decide whether to work with populations to obtain improved versions of elite lines ($F_2$ populations), with sample populations having considerable variation for development of new lines (open-pollinated and composite varieties), or with sample populations that seemingly have less genetic variation (hopefully at a higher frequency) for new lines than other types (synthetic varieties). Ideally, applied breeding programs with short and long-term objectives

---

**Fig. 5.1** Distribution of estimates of additive genetic variance for grain yield for five types of maize populations (see Table 5.2)
should be able to work improving all types of populations to make use of maximum genetic variability for yield and other traits. Fountain and Hallauer (1996) estimated the genetic variability for three types of populations and found that the genetic variability within narrow-base synthetic populations was smaller than within $F_2$ populations; the synthetic varieties were developed by intermating related inbred lines (e.g., B14, Mo17, B73 synthetics) and likely had allele frequencies greater than 0.5.

Some of the first estimates of genetic components of variance in maize were obtained from $F_2$ populations, which were chosen because their expected allele frequencies were 0.5 to provide estimates of average level of dominance of genes affecting the traits studied. One of the first questions raised was the effect of linkage bias on estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. As seen before, if linkages were primarily in the repulsion phase, linkage bias would underestimate $\hat{\sigma}_A^2$ and overestimate $\hat{\sigma}_D^2$. Estimates of average level of dominance for yield often were in the overdominance range, suggesting that at least some loci were giving an expression of overdominance or that pseudo-overdominance was expressed because of linkage bias in estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$. The next logical step was to test for effects of linkage bias on estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ by randomly mating the $F_2$ populations for several generations to permit genetic recombination and an approach to linkage equilibrium. The $F_2$ and advanced random-mated generations ($F_n$) of $F_2$ populations were re-sampled and estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ compared for possible linkage bias. Results from such studies are given in Table 5.3 for four traits.

Estimates of average level of dominance for $F_n$ populations are lower than those for $F_2$ populations in all instances. For yield, the average of estimates for average level of dominance was slightly greater than 1 for the $F_2$ populations, whereas the average of estimates was approximately 0.6 for $F_n$ populations. Average estimates of average level of dominance were about 0.5 for plant and ear height and ear number in $F_2$ populations, but they were about 0.3 in $F_n$ populations. Although estimates were smaller, average levels of dominance in $F_2$ and $F_n$ populations followed the same trend when unbiased averages of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ were used to estimate average levels of dominance. It seems, therefore, that partial dominance rather than complete or overdominance was the predominant expression of gene action for yield. Estimates of overdominance obtained in $F_2$ populations apparently were pseudo-overdominance because of repulsion phase linkages. Levels of dominance were not as great for the other three traits, but linkages also must have influenced estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ because levels of dominance were reduced in all instances. Comparisons of average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ show that $\hat{\sigma}_A^2$ increased and $\hat{\sigma}_D^2$ decreased in all instances from $F_2$ to $F_n$ populations. The last column of Table 5.3 shows the relative change of average level of dominance, which ranges from about 40 to 50 % from $F_2$ to $F_n$ populations.

Compilation of experimental estimates of the parameters shown in Tables 5.1 and 5.2 for 19 different agronomic traits of maize points to one general conclusion:

Genetic variability is present in the five arbitrary types of maize populations for all traits and a major portion of the genetic variability is additive genetic variance. For selection purposes additive genetic variance is of primary importance, and thus
Table 5.3  Estimates of additive (\(\hat{\sigma}^2_A\)) and dominance (\(\hat{\sigma}^2_D\)) variances and ratio of \(\hat{\sigma}^2_D/\hat{\sigma}^2_A\), and average degree of dominance ("a") for four traits estimated in the \(F_2\) populations and \(F_n\) populations advanced by random mating

<table>
<thead>
<tr>
<th>Trait</th>
<th>(F_2)</th>
<th>(F_n)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{\sigma}^2_A)</td>
<td>(\hat{\sigma}^2_D)</td>
<td>(\hat{\sigma}^2_D/\hat{\sigma}^2_A)</td>
</tr>
<tr>
<td>Yield (g)</td>
<td>554.5 (13)</td>
<td>537.5 (13)</td>
<td>1.1021 (0.9693)</td>
</tr>
<tr>
<td>Plant height (cm)</td>
<td>189.4 (7)</td>
<td>42.0 (7)</td>
<td>0.2599 (0.2217)</td>
</tr>
<tr>
<td>Ear height (cm)</td>
<td>84.3 (7)</td>
<td>14.4 (7)</td>
<td>0.2571 (0.1708)</td>
</tr>
<tr>
<td>Ear number ((\times 10^3))</td>
<td>43.5 (8)</td>
<td>10.2 (8)</td>
<td>0.2906 (0.2344)</td>
</tr>
</tbody>
</table>

\(^a\) Numbers in parentheses indicate the number of estimates included in the average
selection should be effective in most populations. The most surprising feature of experimental estimates is lack of any significant trend in relative magnitude of additive genetic variance among five types of populations for yield (see Fig. 5.1). Distributions of estimates of $\hat{\sigma}^2_A$ for the populations themselves do not show any major advantage for one type of population; a priori it seemed that F2 populations would have the least and composites the greatest. However, the summary of estimates was discussed ignoring effects of epistasis and genotype–environment interactions, which will be considered later.

5.2 Iowa Stiff Stalk Synthetic (BSSS)

Iowa Stiff Stalk Synthetic (BSSS) is a synthetic variety developed in 1933–1934 by G. F. Sprague (1946) from the following 16 lines (Table 5.4):

| Set of inbred lines with above average stalk lodging resistance intermated to form BSSS |
|---------------------------------|-----------|-----------------|-----------------|---------------|
| Iowa Stiff Stalk Synthetic (BSSS) |
| Ia.I159 | Ind.461-3 | Ill.Hy | F1B1-7-1 |
| Ia.I224 | Ill.12E | Oh.3167B | A3G-3-1-3 |
| Ia.Os420 | C.I.617 | Ind.AH83 | C.I.187-2 |
| Ia.WD456 | C.I.540 | Ind.Tr9-1-1-6 | LE23 |

They were developed by various breeders and were chosen for their resistance to stalk breakage. BSSS is considered above average as a source population of inbred lines that are above average for combining ability with other elite lines (Mikel, 2006). Surveys by the American Seed Trade Association (ASTA) showed that lines originating from BSSS were used extensively in hybrids in the US Corn Belt (Sprague, 1971; Zuber, 1975). As an example, B14, an elite line extensively used in the northern USA, was derived from BSSS. The most famous line used by industry (see Chapter 1), B73, was derived from an improved version of BSSS. BSSS has been used extensively in selection programs for yield improvement and resistance to maize pests. Relative to other maize populations BSSS is characterized as having above average stalk quality, vigorous plant type, dark green leaf coloration, good ear size, no distinctive features for pest resistance, and full-season maturity for the central US Corn Belt. Yield of BSSS as a variety itself is only average, but it has above average combining ability in crosses with other synthetic varieties (Hallauer and Malithano, 1976). Inbred lines extracted from BSSS have given excellent hybrid yield performance in crosses with Lancaster-type lines (Zuber and Darrah, 1980; Darrah and Zuber, 1986). BSSS seems more uniform in phenotype than other maize populations, but crosses among elite lines extracted from BSSS often express high yields, suggesting genetic variability for yield within BSSS (Lamkey and Hallauer, 1986).
Quantitative genetic studies have been conducted to characterize genetic variability present in the BSSS population. Estimates of genetic parameters for BSSS have been obtained by use of design I and II mating schemes, unselected inbred lines, and recurrent selection programs involving half-sib and inbred progeny evaluations. Detailed estimates of genetic components of variance for BSSS are given as an example of how estimates of a particular population either agree or disagree with those given in Tables 5.1 and 5.2, which, however, include estimates for BSSS.

Table 5.5 shows estimates for eight traits for four sets of experiments. Our purpose is to summarize the estimates of genetic components of variance and to show how they compare by reference to data of studies that used different types of progenies and methods of estimation. Details of specific experiments were reported previously (Hallauer 1970, 1971; Obilana and Hallauer, 1974; Silva and Hallauer, 1975; and Bartual and Hallauer, 1976). In Table 5.5, experiment 1 used progenies developed by design II mating scheme; experiment 3 used progenies developed from design I and II mating schemes; and experiments 2 and 4 included unselected inbred lines developed by single-seed descent by selfing and full-sibbing methods of inbreeding, respectively. Hence estimates of genetic parameters were obtained from evaluation of non-inbred progenies for experiments 1 and 3 and of inbred progenies for experiments 2 and 4. In all instances, however, estimates of genetic components of variance were adjusted for inbreeding and made equivalent for the base population, BSSS.

Comparisons of estimates of components of variance among the four experiments show some large differences, but in most instances the estimates show striking similarities. If we use the standard errors of estimates of components of variance to compare the independent estimates, most are within range of the sampling error. Three of the estimates of $\hat{\sigma}^2_A$ for yield were very similar (experiments 1, 2, and 3, Table 5.5) whereas the estimates of experiment 4 were considerably greater. Experiment 4 estimates were obtained in two similar environments, and consequently the estimate of $\hat{\sigma}^2_{AE}$ was small; the experiment 4 estimate may be overestimated because of bias due to genotype–environment interaction. Estimates of components of genetic variance and their interactions with environments were amazingly similar for experiments 1 and 3 that included the evaluation of non-inbred full-sib progenies; all estimates were similar in all instances. Both estimates of $\hat{\sigma}^2_A$ for plant height obtained from experiments 2 and 4, which included inbred progeny evaluation, were greater than those for experiments 1 and 3, which included non-inbred progeny evaluation. Greater discrepancies existed among estimates for days to flower than for other traits, but the estimates were obtained in fewer environments; if genotype–environment interaction were important, the bias on the estimates would be greater, particularly for inbred progenies. Estimates of heritability were similar for most traits, but they tended to be greater for experiments that evaluated inbred progenies. For yield, estimates of heritability from inbred data were about double of those from non-inbred data. Because only one source of variation was present for inbred progeny evaluation trials, it was not possible to estimate $\hat{\sigma}^2_D$, which would be negligible at the level of inbreeding of the inbred lines used ($F = 99\%$).
Table 5.5  Summary of the component of variance estimates for BSSS obtained from four experiments

<table>
<thead>
<tr>
<th>Trait</th>
<th>Experiment</th>
<th>$\hat{\sigma}^2_A$</th>
<th>$\hat{\sigma}^2_{AE}$</th>
<th>$\hat{\sigma}^2_D$</th>
<th>$\hat{\sigma}^2_{DE}$</th>
<th>$\hat{\sigma}^2_D/\hat{\sigma}^2_A$</th>
<th>$\hat{\sigma}^2$</th>
<th>Plot</th>
<th>Progeny Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yield</td>
<td>1</td>
<td>156 ± 29</td>
<td>83 ± 22</td>
<td>174 ± 37</td>
<td>74 ± 13</td>
<td>1.12</td>
<td>387 ± 13</td>
<td>17.8</td>
<td>34.9</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>147 ± 16</td>
<td>44 ± 5</td>
<td>180 ± 24</td>
<td>71 ± 12</td>
<td>1.10</td>
<td>364 ± 8</td>
<td>18.9</td>
<td>59.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>283 ± 29</td>
<td>22 ± 6</td>
<td></td>
<td></td>
<td></td>
<td>337 ± 10</td>
<td>61.2</td>
<td>89.4</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>188 ± 24</td>
<td>60 ± 11</td>
<td>179 ± 29</td>
<td>73 ± 12</td>
<td>1.11</td>
<td>259 ± 10</td>
<td>24.8</td>
<td>41.4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>143 ± 15</td>
<td>22 ± 4</td>
<td>25 ± 3</td>
<td>5 ± 7</td>
<td>0.17</td>
<td>55 ± 2</td>
<td>57.2</td>
<td>76.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>191 ± 18</td>
<td>15 ± 2</td>
<td></td>
<td></td>
<td></td>
<td>58 ± 3</td>
<td>72.3</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>141 ± 10</td>
<td>12 ± 1</td>
<td>15 ± 2</td>
<td>7 ± 1</td>
<td>0.11</td>
<td>46 ± 1</td>
<td>63.8</td>
<td>82.9</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>191 ± 18</td>
<td>4 ± 2</td>
<td></td>
<td></td>
<td></td>
<td>47 ± 3</td>
<td>78.9</td>
<td>95.4</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>166 ± 15</td>
<td>13 ± 2</td>
<td>20 ± 4</td>
<td>6 ± 4</td>
<td>0.14</td>
<td>52 ± 2</td>
<td>64.6</td>
<td>81.7</td>
</tr>
<tr>
<td>Ear height</td>
<td>1</td>
<td>126 ± 13</td>
<td>10 ± 2</td>
<td>13 ± 3</td>
<td>9 ± 4</td>
<td>0.10</td>
<td>30 ± 1</td>
<td>67.0</td>
<td>83.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>102 ± 10</td>
<td>3 ± 1</td>
<td></td>
<td></td>
<td></td>
<td>35 ± 2</td>
<td>72.8</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>107 ± 7</td>
<td>9 ± 1</td>
<td>9 ± 1</td>
<td>6 ± 1</td>
<td>0.08</td>
<td>31 ± 1</td>
<td>66.0</td>
<td>84.8</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>111 ± 11</td>
<td>1 ± 1</td>
<td></td>
<td></td>
<td></td>
<td>26 ± 2</td>
<td>80.4</td>
<td>96.0</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>112 ± 10</td>
<td>6 ± 1</td>
<td>11 ± 2</td>
<td>8 ± 2</td>
<td>0.09</td>
<td>30 ± 2</td>
<td>67.1</td>
<td>84.4</td>
</tr>
<tr>
<td>Ear length ($\times 10^2$)</td>
<td>1</td>
<td>104 ± 14</td>
<td>22 ± 9</td>
<td>24 ± 12</td>
<td>$-10 \pm 23^d$</td>
<td>0.23</td>
<td>185 ± 6</td>
<td>31.0</td>
<td>62.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>129 ± 15</td>
<td>58 ± 26</td>
<td></td>
<td></td>
<td></td>
<td>242 ± 14</td>
<td>30.1</td>
<td>68.4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>121 ± 11</td>
<td>22 ± 3</td>
<td>41 ± 6</td>
<td>23 ± 4</td>
<td>0.34</td>
<td>121 ± 3</td>
<td>28.9</td>
<td>61.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>215 ± 20</td>
<td>13 ± 5</td>
<td></td>
<td></td>
<td></td>
<td>120 ± 10</td>
<td>61.8</td>
<td>89.8</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>142 ± 15</td>
<td>29 ± 11</td>
<td>32 ± 9</td>
<td>6 ± 14</td>
<td>0.28</td>
<td>167 ± 8</td>
<td>37.8</td>
<td>66.5</td>
</tr>
<tr>
<td>Ear diameter ($\times 10^2$)</td>
<td>1</td>
<td>2.6 ± 0.4</td>
<td>1.1 ± 0.3</td>
<td>0.4 ± 0.4</td>
<td>0.8 ± 0.8</td>
<td>0.15</td>
<td>5.8 ± 0.2</td>
<td>24.3</td>
<td>56.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.1 ± 2.0</td>
<td>3.9 ± 1.5</td>
<td></td>
<td></td>
<td></td>
<td>11.5 ± 1.0</td>
<td>21.0</td>
<td>56.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.9 ± 0.3</td>
<td>0.5 ± 0.1</td>
<td>0.8 ± 0.1</td>
<td>0.7 ± 0.1</td>
<td>0.28</td>
<td>4.7 ± 0.1</td>
<td>30.2</td>
<td>59.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6.0 ± 0.5</td>
<td>0.2 ± 0.1</td>
<td></td>
<td></td>
<td></td>
<td>4.0 ± 0.2</td>
<td>58.8</td>
<td>89.0</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>3.9 ± 0.8</td>
<td>1.4 ± 0.5</td>
<td>0.6 ± 0.2</td>
<td>0.8 ± 0.4</td>
<td>0.22</td>
<td>6.5 ± 0.4</td>
<td>29.5</td>
<td>52.0</td>
</tr>
</tbody>
</table>
Table 5.5 (continued)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Experimenta</th>
<th>$\hat{\sigma}^2_A$</th>
<th>$\hat{\sigma}^2_{AE}$</th>
<th>$\hat{\sigma}^2_D$</th>
<th>$\hat{\sigma}^2_{DE}$</th>
<th>$\hat{\sigma}^2_D/\hat{\sigma}_A^2$</th>
<th>$\hat{\sigma}^2$</th>
<th>Plot</th>
<th>Progeny Meanb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cob diameter ($\times 10^2$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$1.9 \pm 0.2$</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.1 \pm 0.2$</td>
<td>$0.1 \pm 0.4$</td>
<td>$0.05$</td>
<td>$2.8 \pm 0.1$</td>
<td>$38.0$</td>
<td>$75.0$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$1.8 \pm 1.0$</td>
<td>$1.4 \pm 0.0$</td>
<td>$1.4 \pm 0.2$</td>
<td>$1.4 \pm 0.1$</td>
<td>$1.4 \pm 0.2$</td>
<td>$1.4 \pm 0.1$</td>
<td>$2.4 \pm 0.1$</td>
<td>$31.8$</td>
<td>$65.6$</td>
</tr>
<tr>
<td>3</td>
<td>$1.4 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$2.0 \pm 0.1$</td>
<td>$47.5$</td>
<td>$73.7$</td>
</tr>
<tr>
<td>4</td>
<td>$1.9 \pm 0.2$</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.1 \pm 0.1$</td>
<td>$3.3 \pm 0.1$</td>
<td>$30.5$</td>
<td>$65.4$</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>$1.8 \pm 0.4$</td>
<td>$0.4 \pm 0.1$</td>
<td>$0.2 \pm 0.2$</td>
<td>$0.2 \pm 0.2$</td>
<td>$0.2 \pm 0.2$</td>
<td>$0.2 \pm 0.2$</td>
<td>$4.5 \pm 0.2$</td>
<td>$19.2$</td>
<td>$47.9$</td>
</tr>
<tr>
<td>Kernel depth ($\times 10^2$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$1.5 \pm 0.3$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.5 \pm 0.3$</td>
<td>$0.5 \pm 0.3$</td>
<td>$0.5 \pm 0.3$</td>
<td>$0.5 \pm 0.3$</td>
<td>$4.8 \pm 0.2$</td>
<td>$19.2$</td>
<td>$47.9$</td>
</tr>
<tr>
<td>2</td>
<td>$1.4 \pm 0.0$</td>
<td>$1.0 \pm 0.0$</td>
<td>$1.0 \pm 0.0$</td>
<td>$1.0 \pm 0.0$</td>
<td>$1.0 \pm 0.0$</td>
<td>$1.0 \pm 0.0$</td>
<td>$5.6 \pm 0.1$</td>
<td>$17.5$</td>
<td>$52.5$</td>
</tr>
<tr>
<td>3</td>
<td>$1.5 \pm 0.2$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$0.2 \pm 0.1$</td>
<td>$3.9 \pm 0.1$</td>
<td>$23.4$</td>
<td>$43.3$</td>
</tr>
<tr>
<td>4</td>
<td>$3.2 \pm 0.4$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$5.6 \pm 0.1$</td>
<td>$35.2$</td>
<td>$75.6$</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>$1.9 \pm 0.5$</td>
<td>$0.6 \pm 0.1$</td>
<td>$0.6 \pm 0.1$</td>
<td>$0.6 \pm 0.1$</td>
<td>$0.6 \pm 0.1$</td>
<td>$0.6 \pm 0.1$</td>
<td>$5.0 \pm 0.1$</td>
<td>$23.4$</td>
<td>$54.8$</td>
</tr>
<tr>
<td>Days to flower</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$30.5 \pm 5.8$</td>
<td>$---$</td>
<td>$---$</td>
<td>$---$</td>
<td>$---$</td>
<td>$---$</td>
<td>$2.2 \pm 0.2$</td>
<td>$74.0$</td>
<td>$77.5$</td>
</tr>
<tr>
<td>2</td>
<td>$8.6 \pm 0.6$</td>
<td>$0.4 \pm 0.2$</td>
<td>$0.4 \pm 0.2$</td>
<td>$0.4 \pm 0.2$</td>
<td>$0.4 \pm 0.2$</td>
<td>$0.4 \pm 0.2$</td>
<td>$4.8 \pm 0.2$</td>
<td>$62.3$</td>
<td>$90.2$</td>
</tr>
<tr>
<td>3</td>
<td>$4.2 \pm 0.3$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$0.3 \pm 0.1$</td>
<td>$1.3 \pm 0.1$</td>
<td>$64.6$</td>
<td>$82.6$</td>
</tr>
<tr>
<td>4</td>
<td>$16.8 \pm 1.6$</td>
<td>$8.6 \pm 0.6$</td>
<td>$8.6 \pm 0.6$</td>
<td>$8.6 \pm 0.6$</td>
<td>$8.6 \pm 0.6$</td>
<td>$8.6 \pm 0.6$</td>
<td>$3.0 \pm 0.3$</td>
<td>$59.2$</td>
<td>$83.3$</td>
</tr>
<tr>
<td>$\bar{X}$</td>
<td>$15.0 \pm 2.1$</td>
<td>$0.4 \pm 0.2$</td>
<td>$4.5 \pm 3.4$</td>
<td>$4.5 \pm 3.4$</td>
<td>$4.5 \pm 3.4$</td>
<td>$4.5 \pm 3.4$</td>
<td>$2.8 \pm 0.2$</td>
<td>$65.5$</td>
<td>$74.4$</td>
</tr>
</tbody>
</table>

a The mating designs used to obtain the component of variance estimates were (1) design II with 320 full-sib progenies evaluated in three environments (Hallauer, 1971); (2) 247 unselected inbred lines developed by selfing evaluated in three environments (Obilana and Hallauer, 1974); (3) 800 full-sib progenies developed by use of designs I and II evaluated in six environments (Silva, 1974); and (4) 231 unselected inbred lines developed by full-sibbing evaluated in two experiments (Bartual and Hallauer, 1976)

b Progeny mean heritability estimates were calculated assuming three environments and two replications in each environment; i.e., $\hat{\sigma}^2_A / (\hat{\sigma}^2_A + \hat{\sigma}^2_e / e + \hat{\sigma}^2_A)$

c No estimates were available

d Negative estimates were assumed to be zero in the calculation of heritability estimates

e Estimates of heritability were computed from the average estimates of the components of variance
Comparison of estimates given in Table 5.1 with those given in Table 5.5 for BSSS shows that estimates for BSSS tended to be smaller than those summarized across all populations. For example, the average estimate of $\hat{\sigma}_A^2$ in Table 5.1 is $469.1 \pm 174.3$ vs. $188 \pm 24$ for BSSS in Table 5.5. The estimate of $\hat{\sigma}_D^2$ (287 $\pm$ 210 in Table 5.1) also is much greater than the estimate of $\hat{\sigma}_D^2$ for BSSS (179 $\pm$ 29 in Table 5.5). Although average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ for BSSS are 60% and 38% smaller than those in Table 5.1, standard errors of estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ also are 84% smaller for BSSS. If the precision of estimates is considered and we use two standard errors of the respective estimates to determine significance, estimates of BSSS are not significantly different from those in Table 5.1 for yield. Estimates of plant height, ear height, ear length, and ear diameter for BSSS are similar to those given in Table 5.1 and are within the range of sampling error. However, variation of cob diameter and kernel depth in BSSS is less than that for other populations summarized in Table 5.1; and the estimate of $\hat{\sigma}_A^2$ for days to flower was about four times greater than the average estimate of all populations in Table 5.1. Estimates of $\hat{\sigma}_A^2$ for days to flower for BSSS are probably biased upward because of the genotype–environment interaction. Experiment 3 in Table 5.5 had an estimate of $\hat{\sigma}_A^2$ for days to flower that was similar to that given in Table 5.1.

Heritability estimates on a plot basis were very similar for BSSS and the average of all populations for yield, plant height, ear height, ear length, ear diameter, cob diameter, kernel depth, and days to flower. Considering sampling errors associated with estimates of components of variance, the similarities of heritability estimates in Tables 5.1 and 5.5 are remarkable. Relative heritability estimates for the eight traits measured in BSSS are the same as those for the average of all populations in Table 5.1.

The most distinctive difference between the BSSS estimates and the average for all varieties was in the estimate of $\hat{\sigma}_D^2$ for yield. The average estimate of $\hat{\sigma}_D^2$ across populations is 39% smaller than the average estimate of $\hat{\sigma}_A^2$, whereas average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ for BSSS are nearly equal (Table 5.5). The averages of the ratio $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ are about one in both instances (0.94 vs. 1.11), but ratios of average estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ are 0.61 (Table 5.1) vs. 0.95 (Table 5.5). If, for the BSSS population, we consider only experiments 1 and 3 for which estimates of both $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ were obtained, the average estimate of $\hat{\sigma}_A^2$ (161) is less than the average estimate of $\hat{\sigma}_A^2$ (179); hence $\hat{\sigma}_D^2/\hat{\sigma}_A^2$ is 1.18. In addition, comparison of estimates of $\hat{\sigma}_A^2$ of BSSS with $\hat{\sigma}_A^2$ in Table 5.2 for the average of 15 studies of synthetic varieties shows that the estimates are similar, 188 vs. 226. The average estimate of $\hat{\sigma}_D^2$ for BSSS, however, is greater than the average estimate for all synthetic varieties, albeit the standard errors show that the difference between the two average estimates of $\hat{\sigma}_D^2$ is probably not significant. It seems that the BSSS population relative to other synthetic varieties is average for $\hat{\sigma}_A^2$ and above average for $\hat{\sigma}_D^2$ in the case of grain yield. For the other traits, BSSS has genetic variability similar to other synthetic varieties for plant height, ear height, ear length, and ear diameter but below average for cob diameter and kernel depth. Except for $\hat{\sigma}_D^2$, there is no evidence to distinguish
BSSS from other synthetic variety populations. The estimates of $\hat{\sigma}_A^2$ and $\hat{h}^2$ were similar to the average of all synthetic variety populations. Compared to the mean of all populations estimates of $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ were smaller, but standard errors of estimates also were smaller. Resistance to stalk breakage is a very important trait in the production and growing of hybrids in mechanized maize culture. Perhaps the fortuitous choice of lines with good stalk strength and strong dominance of genes for yield were the contributing factors for BSSS being a good source population of inbred lines having good combining ability (Hallauer et al., 1983; Mikel, 2006; Mikel and Dudley, 2006).

### 5.3 Selection Experiments vs. Mating Designs for Prediction

Long-term selection programs for the cyclical improvement of breeding populations as source materials for applied breeding programs also provide evidence of the genetic variability present within populations. Estimates of genetic variability are only as good as the sampling techniques used in obtaining progenies for evaluation. Effects of sampling are equally important whether one is sampling populations to develop progenies by some mating design for estimation of variance components or sampling populations to test progenies in a cyclical improvement program. In either instance, the validity of estimates of variance components and progeny evaluations for populations is only as good as the sampling used. In most instances the sample sizes were greater when mating designs (e.g., designs I and II) were used than when testing progenies included in evaluation trials of cyclical selection programs. In cyclical selection programs, however, estimates of genetic variation among progenies tested in each cycle of selection are available. As shown in Chapter 2, components of genetic variance are dependent on allele frequencies of the population under study; changes in allele frequency of the population would affect estimates of components of variance. The primary objective of all selection trials is to change allele frequency, i.e., to increase the frequency of desirable alleles for the trait under selection. Use of variation among progenies after several cycles of selection as an estimate of the component of variance for original unselected populations, therefore, may be invalid. If, as expected, changes in allele frequency are relatively small from cycle to cycle of selection for a complex trait such as yield, the small change would not have a large effect on estimates of components of variance. Sampling errors may in fact be greater than absolute changes in allele frequency in the precision of estimates of components of variance. Estimates of components of variance obtained by combining data from long-term cyclical selection programs may provide better estimates of the population under selection because of repeated sampling of the population in each cycle of selection, assuming changes in allele frequency are not great in successive cycles of selection. Sufficient sampling and change in allele frequency, therefore, are antagonistic in regard to estimates from selection experiments compared with a study by use of a mating design.
### Table 5.6

Populations undergoing recurrent selection for greater grain yield, greater root and stalk strength, improved tolerance to foliar leaf diseases, and earlier maturity in the Iowa State University (ISU) maize breeding program at Ames, Iowa

<table>
<thead>
<tr>
<th>Selection Population</th>
<th>Designation</th>
<th>Tester</th>
<th>Progenies evaluated</th>
<th>Cycles completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iowa Stiff Stalk Synthetic</td>
<td>BS13(HT)</td>
<td>Ia13</td>
<td>Half-sibs</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>BS13(S)</td>
<td>—</td>
<td>S₂ progenies</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>BS13(HI)</td>
<td>B97</td>
<td>Half-sibs</td>
<td>6</td>
</tr>
<tr>
<td>Iowa Stiff Stalk Synthetic</td>
<td>BSSS(R)</td>
<td>BSCB1</td>
<td>Half-sibs/Full-sibs</td>
<td>18</td>
</tr>
<tr>
<td>Iowa Corn Borer Synthetic</td>
<td>BSCB1(R)</td>
<td>BSSS</td>
<td>Half-sibs/Full-sibs</td>
<td>18</td>
</tr>
<tr>
<td>Synthetic No. 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krug Hi I. Synthetic 3</td>
<td>BSK(HI)</td>
<td>Inbred</td>
<td>Half-sibs</td>
<td>11</td>
</tr>
<tr>
<td>Krug Hi I. Synthetic 3</td>
<td>BSK(S)</td>
<td>—</td>
<td>S₁–S₂ progenies</td>
<td>11</td>
</tr>
<tr>
<td>Iowa Two-ear Synthetic</td>
<td>BS10(FR)</td>
<td>BS11</td>
<td>Full-sibs</td>
<td>18</td>
</tr>
<tr>
<td>Pioneer Two-ear Synthetic</td>
<td>BS11(FR)</td>
<td>BS10</td>
<td>Full-sibs</td>
<td>18</td>
</tr>
<tr>
<td>Pioneer Two-ear Synthetic</td>
<td>BS11(S)</td>
<td>—</td>
<td>S₁ progenies</td>
<td>11</td>
</tr>
<tr>
<td>Syntheticc</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iowa Early Synthetic No. 1</td>
<td>BS21(R)</td>
<td>BS22</td>
<td>Half-sibs</td>
<td>9</td>
</tr>
<tr>
<td>Iowa Early Synthetic No. 2</td>
<td>BS22(R)</td>
<td>BS21</td>
<td>Half-sibs</td>
<td>9</td>
</tr>
<tr>
<td>Tuxpeno Composited</td>
<td>BS28(R)</td>
<td>BS29</td>
<td>Half-sibs</td>
<td>5</td>
</tr>
<tr>
<td>Suwan-1d</td>
<td>BS29(R)</td>
<td>BS28</td>
<td>Half-sibs</td>
<td>5</td>
</tr>
<tr>
<td>Midland</td>
<td>BS33(S)</td>
<td>—</td>
<td>S₁–S₂ progenies</td>
<td>4</td>
</tr>
<tr>
<td>Lancaster Synthetic</td>
<td>BS34(S)</td>
<td>—</td>
<td>S₁–S₂ progenies</td>
<td>4</td>
</tr>
</tbody>
</table>

---

*a BS13(HI) – selection with B97 as tester was initiated in BS13(HT)C7, which is the same source population for BS13(S)*

*b Selection based on half-sibs was continued through C12, after which selection was based on full-sibs*

*c Includes four selection programs where 5, 10, 20, and 30 selections are inter-mated after each cycle of selection*

*d Mass selection was used to adapt two tropical populations to temperate environments before initiation of RRS*

Table 5.6 shows all the recurrent selection studies that were being conducted at the Iowa station (Ames, IA) at the time the state corn breeding position was active. Few public corn breeding programs are left conducting programs that link recurrent selection for germplasm improvement with inbred line development (see Chapter 7 for ongoing programs at the Fargo station).

Estimates of progeny components of variance and their interactions with environments were summarized for the following selection studies; some programs were discontinued, while others were emphasized:
Hereditary Variance: Experimental Estimates

(1) BS13(HT): Iowa Stiff Stalk Synthetic with seven cycles of half-sib recurrent selection with tester (H = half-sibs, T = tester)
(2) BS13(S): Iowa Stiff Stalk Synthetic with four cycles of inbred progeny recurrent selection (S = either S1 and/or S2 recurrent selection)
(3) BSSS(R): Iowa Stiff Stalk Synthetic with 10 cycles of reciprocal recurrent selection with Iowa Corn Borer Synthetic No. 1 (BSCB1) as the tester population (R = reciprocal recurrent selection based on half-sib progenies)
(4) BSCB1(R): Iowa Corn Borer Synthetic No. 1 with 10 cycles of reciprocal recurrent selection with Iowa Stiff Stalk Synthetic as the tester population
(5) BSK(HI): Krug Hi Synthetic 3 (BSK), a strain of Krug Yellow Dent (Lonnquist, 1949), with eight cycles of half-sib recurrent selection with inbred line testers (HI = half-sibs)
(6) BSK(S): Krug Hi Synthetic 3 with eight cycles of S1 progeny recurrent selection
(7) BS12(HI): Alph, an open-pollinated variety with seven cycles of half-sib recurrent selection with B14 line as tester
(8) BS2(S), BS16(S), and (8) BSTL(S) exotic populations that were undergoing S2 recurrent selection, and
(9) BS10(FR) and BS11(FR) undergoing reciprocal full-sib selection (FR = reciprocal selection based on full-sib progenies). Grain yield was emphasized in all instances.

All of the above selection programs in Table 5.6 were undergoing, except for (7) and (8). The purposes of providing estimates of genetic variation for recurrent selection studies are two-fold:

(1) To compare estimates obtained from selection trials with those shown in Tables 5.1–5.4 obtained via different mating designs; and
(2) To show how estimates of genetic variance have changed in successive cycles of recurrent selection.

Some studies have been in progress for nearly 40–60 years, others for only 20. Experimental and testing procedures have changed during the selection studies and these will be indicated for each. All estimates of $\hat{\sigma}_A^2$ were converted to grams per plant to make them equivalent to those in Tables 5.1–5.4. In making the conversions, we used an average plant density of 39,042 plants/ha, the plant density used in the earlier years; current testing plant densities range from 60,000 to 70,000 plants/ha.

Detailed data of the experiments that have undergone recurrent selection for seven or more cycles are summarized in Tables 5.7–5.12. Estimates of variance components, heritability, and genetic coefficients of variation for each cycle of selection are included.

In all cases, there is a similar trend for each population for successive cycles of selection: Estimates of the progeny component of variance ($\hat{\sigma}_g^2$) are greatest in the initial two cycles of selection, are smallest for cycles 2–4, and increase in the following cycles. This trend is not as great for the BSK(S) selection population as...
### Table 5.7 Summary of the estimates of variance components for grain yield for BS13(HT) with Ia13 as the tester for seven cycles of half-sib recurrent selection and two cycles of S<sup>2</sup> recurrent selection

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Year</th>
<th>No. of trials</th>
<th>Reps/trial</th>
<th>Variance component estimates&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Genetic yield&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Average yield&lt;sup&gt;c&lt;/sup&gt; (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS13(HT)C0</td>
<td>1940</td>
<td>1</td>
<td>3</td>
<td>41.9 ± 3.2</td>
<td>12.9 ± 3.3</td>
<td>338.5 ± 86.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>338.5 ± 86.6</td>
<td>48.0</td>
<td>10.2</td>
</tr>
<tr>
<td>BS13(HT)C1</td>
<td>1948</td>
<td>1</td>
<td>6</td>
<td>15.3 ± 0.8</td>
<td>4.2 ± 0.8</td>
<td>110.2 ± 21.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110.2 ± 21.0</td>
<td>62.2</td>
<td>3.8</td>
</tr>
<tr>
<td>BS13(HT)C2</td>
<td>1952</td>
<td>2</td>
<td>3</td>
<td>29.3 ± 2.1</td>
<td>9.9 ± 2.6</td>
<td>259.8 ± 68.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>259.8 ± 68.2</td>
<td>58.9</td>
<td>4.8</td>
</tr>
<tr>
<td>BS13(HT)C3</td>
<td>1955</td>
<td>2</td>
<td>3</td>
<td>63.2 ± 5.1</td>
<td>-1.6 ± 3.6</td>
<td>-42.0 ± 94.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-42.0 ± 94.5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>BS13(HT)C4</td>
<td>1958-59</td>
<td>4</td>
<td>3</td>
<td>37.7 ± 2.0</td>
<td>7.0 ± 1.8</td>
<td>89.2 ± 34.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>89.2 ± 34.1</td>
<td>58.6</td>
<td>2.6</td>
</tr>
<tr>
<td>BS13(HT)C5</td>
<td>1962</td>
<td>4</td>
<td>2</td>
<td>32.5 ± 2.5</td>
<td>10.3 ± 2.3</td>
<td>270.3 ± 60.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>270.3 ± 60.4</td>
<td>65.2</td>
<td>3.8</td>
</tr>
<tr>
<td>BS13(HT)C6</td>
<td>1965</td>
<td>4</td>
<td>2</td>
<td>30.4 ± 2.4</td>
<td>9.1 ± 2.2</td>
<td>238.8 ± 57.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>238.8 ± 57.7</td>
<td>56.2</td>
<td>4.5</td>
</tr>
<tr>
<td>Pooled (C0–C6)</td>
<td></td>
<td>1.9&lt;sup&gt;d&lt;/sup&gt;</td>
<td>2.8&lt;sup&gt;d&lt;/sup&gt;</td>
<td>38.0 ± 1.2</td>
<td>6.6 ± 1.0</td>
<td>207.3 ± 28.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>207.3 ± 28.9</td>
<td>42.7</td>
<td>5.0</td>
</tr>
<tr>
<td>BS13(S)C0</td>
<td>1972</td>
<td>3</td>
<td>2</td>
<td>37.2 ± 2.4</td>
<td>23.4 ± 4.5</td>
<td>257.6 ± 45.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>257.6 ± 45.5</td>
<td>76.7</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>257.6 ± 45.5</td>
<td>76.7</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350.5 ± 75.2</td>
<td>62.6</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350.5 ± 75.2</td>
<td>62.6</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>199.3 ± 99.6</td>
<td>69.2</td>
<td>12.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>199.3 ± 99.6</td>
<td>69.2</td>
<td>12.9</td>
</tr>
<tr>
<td>BS13(S)C1</td>
<td>1975</td>
<td>3</td>
<td>2</td>
<td>125.0 ± 6.1</td>
<td>54.0 ± 12.4</td>
<td>80.2 ± 17.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80.2 ± 17.2</td>
<td>20.0 ± 4.3</td>
<td>350.5 ± 75.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350.5 ± 75.2</td>
<td>62.6</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350.5 ± 75.2</td>
<td>62.6</td>
<td>22.3</td>
</tr>
<tr>
<td>BS13(S)C2</td>
<td>1978</td>
<td>3</td>
<td>2</td>
<td>78.4 ± 6.6</td>
<td>21.5 ± 6.9</td>
<td>45.6 ± 22.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45.6 ± 22.8</td>
<td>199.3 ± 99.6</td>
<td>69.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45.6 ± 22.8</td>
<td>199.3 ± 99.6</td>
<td>69.2</td>
</tr>
<tr>
<td>BS13(S)C3</td>
<td>1981</td>
<td>3</td>
<td>2</td>
<td>65.8 ± 6.7</td>
<td>48.4 ± 9.6</td>
<td>27.8 ± 8.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.8 ± 8.4</td>
<td>121.5 ± 36.7</td>
<td>50.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.8 ± 8.4</td>
<td>121.5 ± 36.7</td>
<td>50.7</td>
</tr>
<tr>
<td>BS13(S)C4</td>
<td>1984</td>
<td>3</td>
<td>2</td>
<td>42.8 ± 3.6</td>
<td>10.5 ± 3.7</td>
<td>28.4 ± 5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.4 ± 5.7</td>
<td>124.1 ± 24.9</td>
<td>72.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.4 ± 5.7</td>
<td>124.1 ± 24.9</td>
<td>72.8</td>
</tr>
<tr>
<td>Pooled (C0–C4)</td>
<td></td>
<td>3</td>
<td>2</td>
<td>65.4 ± 2.5</td>
<td>32.0 ± 3.2</td>
<td>234.7 ± 52.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>234.7 ± 52.4</td>
<td>71.3</td>
<td>15.9</td>
</tr>
</tbody>
</table>

<sup>a</sup> Variance component estimates in quintals per hectare

<sup>b</sup> Estimates of \( \hat{\sigma}^2 \) for half-sib and S<sup>2</sup> selection converted to grams per plant

<sup>c</sup> Values adjusted to \((\frac{1}{4})\hat{\sigma}^2\) relative to half-sib selection

<sup>d</sup> Values are harmonic means
Table 5.8 Summary of the estimates of variance components for grain yield for BSSS(R) with BSCB1(R) as the tester in a reciprocal recurrent selection experiment

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Year</th>
<th>No. of trials</th>
<th>Reps/trial</th>
<th>( \hat{\sigma}^2 )</th>
<th>( \hat{\sigma}_{ge}^2 )</th>
<th>( \hat{\sigma}_g^2 )</th>
<th>( \hat{\sigma}_{A}^2 )</th>
<th>( h^2 )</th>
<th>Genetic CV (%)</th>
<th>Average yield (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS(R)C0</td>
<td>1950</td>
<td>1</td>
<td>3</td>
<td>27.9 ± 3.0</td>
<td>—</td>
<td>13.3 ± 3.4</td>
<td>349.0 ± 89.2</td>
<td>58.8</td>
<td>7.9</td>
<td>46.2</td>
</tr>
<tr>
<td>BSSS(R)C1</td>
<td>1953</td>
<td>2</td>
<td>3</td>
<td>33.1 ± 2.3</td>
<td>0.1 ± 1.7</td>
<td>11.1 ± 2.5</td>
<td>291.3 ± 65.6</td>
<td>66.7</td>
<td>6.2</td>
<td>53.4</td>
</tr>
<tr>
<td>BSSS(R)C2</td>
<td>1956–57</td>
<td>4</td>
<td>2.7(^c)</td>
<td>41.4 ± 2.2</td>
<td>3.3 ± 1.7</td>
<td>3.9 ± 1.2</td>
<td>102.3 ± 31.5</td>
<td>43.5</td>
<td>3.7</td>
<td>53.2</td>
</tr>
<tr>
<td>BSSS(R)C3</td>
<td>1960</td>
<td>2</td>
<td>3</td>
<td>24.5 ± 1.7</td>
<td>4.0 ± 1.8</td>
<td>4.2 ± 1.7</td>
<td>110.2 ± 44.6</td>
<td>43.7</td>
<td>2.8</td>
<td>74.5</td>
</tr>
<tr>
<td>BSSS(R)C4</td>
<td>1964</td>
<td>4</td>
<td>2</td>
<td>30.2 ± 2.4</td>
<td>3.0 ± 2.0</td>
<td>2.0 ± 1.0</td>
<td>52.5 ± 26.2</td>
<td>27.5</td>
<td>2.1</td>
<td>66.9</td>
</tr>
<tr>
<td>BSSS(R)C5</td>
<td>1970</td>
<td>4</td>
<td>2</td>
<td>64.6 ± 4.1</td>
<td>4.6 ± 4.7</td>
<td>13.2 ± 3.3</td>
<td>346.4 ± 86.6</td>
<td>56.0</td>
<td>5.3</td>
<td>68.0</td>
</tr>
<tr>
<td>BSSS(R)C6</td>
<td>1973</td>
<td>3</td>
<td>2</td>
<td>70.6 ± 4.5</td>
<td>14.0 ± 5.9</td>
<td>15.6 ± 5.5</td>
<td>409.3 ± 144.3</td>
<td>45.4</td>
<td>5.5</td>
<td>71.3</td>
</tr>
<tr>
<td>BSSS(R)C9(^c)</td>
<td>1982</td>
<td>4</td>
<td>2</td>
<td>104.0 ± 5.7</td>
<td>8.0 ± 5.6</td>
<td>17.2 ± 3.8</td>
<td>357.4 ± 79.0</td>
<td>53.4</td>
<td>4.7</td>
<td>88.2</td>
</tr>
<tr>
<td>BSSS(R)C10(^c)</td>
<td>1985</td>
<td>4</td>
<td>2</td>
<td>45.4 ± 3.1</td>
<td>5.8 ± 2.7</td>
<td>5.7 ± 1.8</td>
<td>118.4 ± 37.4</td>
<td>45.0</td>
<td>3.3</td>
<td>73.5</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>2.6(^d)</td>
<td>2.3(^d)</td>
<td>57.8 ± 3.8</td>
<td>7.8 ± 4.1</td>
<td>14.1 ± 3.6</td>
<td>359.3 ± 92.4</td>
<td>52.8</td>
<td>5.4</td>
<td>69.1</td>
</tr>
</tbody>
</table>

\(^a\) Variance component estimates in quintals per hectare
\(^b\) Estimates of \( \hat{\sigma}_A^2 \) from half-sib and full-sib selection in grams per plant
\(^c\) Estimates for reciprocal full-sib selection with BSCB1 (see Table 5.9)
\(^d\) Values are harmonic means
Table 5.9  Summary of estimates of variance components for grain yield for BSCB1 (R) with BSSS(R) as the tester in reciprocal recurrent selection experiment

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Year</th>
<th>No. of trials</th>
<th>Reps/trial</th>
<th>( \hat{\sigma}^2 )</th>
<th>( \hat{\sigma}^2_{ge} )</th>
<th>( \hat{\sigma}^2_g )</th>
<th>( \hat{\sigma}^2_b )</th>
<th>( \hat{h}^2 )</th>
<th>Genetic CV (%)</th>
<th>Average yield (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSCB1(R)C0</td>
<td>1950</td>
<td>1</td>
<td>3</td>
<td>14.9 ± 1.6</td>
<td>—</td>
<td>20.8 ± 3.7</td>
<td>545.8 ± 97.1</td>
<td>80.7</td>
<td>10.8</td>
<td>42.0</td>
</tr>
<tr>
<td>BSCB1(R)C1</td>
<td>1953</td>
<td>2</td>
<td>3</td>
<td>34.1 ± 2.8</td>
<td>2.3 ± 2.3</td>
<td>16.9 ± 3.8</td>
<td>443.4 ± 99.7</td>
<td>71.2</td>
<td>7.7</td>
<td>53.4</td>
</tr>
<tr>
<td>BSCB1(R)C2</td>
<td>1956–57</td>
<td>4</td>
<td>3</td>
<td>43.1 ± 2.2</td>
<td>4.6 ± 1.7</td>
<td>2.8 ± 1.1</td>
<td>73.5 ± 28.9</td>
<td>37.1</td>
<td>3.2</td>
<td>51.8</td>
</tr>
<tr>
<td>BSCB1(R)C3</td>
<td>1960</td>
<td>2</td>
<td>3</td>
<td>26.0 ± 1.9</td>
<td>7.5 ± 2.4</td>
<td>6.8 ± 2.4</td>
<td>178.4 ± 63.0</td>
<td>45.7</td>
<td>3.5</td>
<td>74.1</td>
</tr>
<tr>
<td>BSCB1(R)C4</td>
<td>1964</td>
<td>4</td>
<td>2</td>
<td>23.5 ± 1.8</td>
<td>5.8 ± 1.8</td>
<td>5.7 ± 1.5</td>
<td>149.6 ± 39.4</td>
<td>56.7</td>
<td>3.6</td>
<td>66.5</td>
</tr>
<tr>
<td>BSCB1(R)C5</td>
<td>1970</td>
<td>4</td>
<td>2</td>
<td>87.6 ± 5.6</td>
<td>12.1 ± 6.8</td>
<td>11.0 ± 3.8</td>
<td>288.6 ± 99.7</td>
<td>44.0</td>
<td>4.9</td>
<td>67.5</td>
</tr>
<tr>
<td>BSCB1(R)C6</td>
<td>1973</td>
<td>3</td>
<td>2</td>
<td>81.3 ± 5.2</td>
<td>11.6 ± 6.4</td>
<td>32.9 ± 8.8</td>
<td>863.3 ± 230.9</td>
<td>65.3</td>
<td>8.0</td>
<td>72.0</td>
</tr>
<tr>
<td>BSCB1(R)C9(^c)</td>
<td>1982</td>
<td>4</td>
<td>2</td>
<td>104.0 ± 5.7</td>
<td>8.0 ± 5.6</td>
<td>17.2 ± 3.8</td>
<td>357.4 ± 79.0</td>
<td>53.4</td>
<td>4.7</td>
<td>88.2</td>
</tr>
<tr>
<td>BSCB1(R)C10(^c)</td>
<td>1985</td>
<td>4</td>
<td>2</td>
<td>45.4 ± 3.1</td>
<td>5.8 ± 2.7</td>
<td>5.7 ± 1.8</td>
<td>118.4 ± 37.4</td>
<td>45.0</td>
<td>3.3</td>
<td>73.5</td>
</tr>
<tr>
<td>Mean</td>
<td>2.6(^d)</td>
<td>2.3(^d)</td>
<td></td>
<td>59.1 ± 3.9</td>
<td>8.2 ± 4.4</td>
<td>13.7 ± 3.7</td>
<td>349.1 ± 93.6</td>
<td>53.8</td>
<td>5.4</td>
<td>68.5</td>
</tr>
</tbody>
</table>

\(^a\) Variance component estimates in quintals per hectare
\(^b\) Estimates of \( \hat{\sigma}^2_b \) from half-sib and full-sib selection in grams per plant
\(^c\) Estimates for reciprocal full-sib selection with BSSS (see Table 5.8)
\(^d\) Values are harmonic means
### Table 5.10 Summary of estimates of variance components for yield for BS12(HI) for seven cycles of half sib recurrent selection by use of the inbred line tester: B14

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Year</th>
<th>No. of trials</th>
<th>Reps/trial</th>
<th>Variance component estimates$^a$</th>
<th>Genetic CV (%)</th>
<th>Average yield (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS12(HI)C0</td>
<td>1950</td>
<td>1</td>
<td>3</td>
<td>$\hat{\sigma}^2$</td>
<td>4.8 ± 0.6</td>
<td>18.8 ± 3.2</td>
</tr>
<tr>
<td>BS12(HI)C1</td>
<td>1954</td>
<td>2</td>
<td>3</td>
<td>$\hat{\sigma}^2_{ge}$</td>
<td>26.8 ± 2.0</td>
<td>15.9 ± 4.5</td>
</tr>
<tr>
<td>BS12(HI)C2</td>
<td>1959</td>
<td>4</td>
<td>3</td>
<td>$\hat{\sigma}^2_{g}$</td>
<td>36.8 ± 1.8</td>
<td>4.4 ± 1.5</td>
</tr>
<tr>
<td>BS12(HI)C3</td>
<td>1963</td>
<td>4</td>
<td>2</td>
<td>$\hat{\sigma}^2_{A}$</td>
<td>28.6 ± 2.2</td>
<td>10.1 ± 2.4</td>
</tr>
<tr>
<td>BS12(HI)C4</td>
<td>1967</td>
<td>3</td>
<td>2</td>
<td>$\hat{h}^2$</td>
<td>46.2 ± 4.2</td>
<td>3.8 ± 1.9</td>
</tr>
<tr>
<td>BS12(HI)C5</td>
<td>1972</td>
<td>4</td>
<td>2</td>
<td>$\hat{\sigma}^2_{A}$</td>
<td>85.5 ± 6.7</td>
<td>19.7 ± 5.3</td>
</tr>
<tr>
<td>BS12(HI)C6</td>
<td>1975</td>
<td>3</td>
<td>2</td>
<td>$\hat{\sigma}^2_{A}$</td>
<td>211.9 ± 17.3</td>
<td>38.0 ± 11.4</td>
</tr>
<tr>
<td>BS12(HI)C7</td>
<td>1979</td>
<td>3</td>
<td>2</td>
<td>$\hat{\sigma}^2_{A}$</td>
<td>141.7 ± 12.9</td>
<td>121.9 ± 8.6</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>2.4$^c$</td>
<td>2.3$^c$</td>
<td>$\hat{\sigma}^2_{A}$</td>
<td>72.8 ± 6.0</td>
<td>14.7 ± 6.4</td>
</tr>
</tbody>
</table>

$^a$ Variance component estimates from the selection experiments in quintals per hectare

$^b$ Estimates of $\hat{\sigma}^2_{A}$ calculated from half-sib selection, converted to grams per plant

$^c$ Values are harmonic means
Table 5.11  Summary of estimates of variance components for grain yield for BSK(HI) for eight cycles of half-sib recurrent selection by use of a tester

| Selection population | Year    | No. of trials | Reps/trial | Variance component estimates\(^\text{a}\) | \(\hat{\sigma}^2\) \(\hat{\sigma}^2_{ge}\) \(\hat{\sigma}^2_g\) \(\hat{\sigma}^2_A\) | \(\hat{h}^2\) | Genetic CV (%) | Average yield (q/ha) |
|----------------------|---------|---------------|------------|---------------------------------------------|-----------------|-----------------|-------------------|
| BSK(HI)C0            | 1954    | 2             | 3          | 25.0 ± 2.8 5.1 ± 2.1 6.4 ± 2.1             | 167.9 ± 55.1    | 48.8            | 4.0               | 63.1             |
| BSK(HI)C1            | 1958–59 | 4             | 2          | 23.1 ± 1.7 6.1 ± 1.6 13.0 ± 2.4             | 341.1 ± 63.0    | 74.7            | 5.2               | 69.4             |
| BSK(HI)C2            | 1962    | 4             | 2          | 26.6 ± 2.1 12.6 ± 2.5 3.0 ± 1.5             | 78.7 ± 39.4     | 31.6            | 2.4               | 71.6             |
| BSK(HI)C3            | 1965    | 4             | 2          | 32.9 ± 2.6 3.8 ± 2.2 3.6 ± 1.4             | 94.5 ± 36.7     | 41.5            | 3.4               | 55.7             |
| BSK(HI)C4            | 1968    | 4             | 2          | 18.5 ± 1.4 6.3 ± 1.5 5.6 ± 1.4             | 146.9 ± 36.7    | 58.9            | 3.4               | 69.2             |
| BSK(HI)C5            | 1971    | 4             | 2          | 76.4 ± 6.0 6.0 ± 4.7 9.9 ± 3.1             | 259.8 ± 81.3    | 57.2            | 5.0               | 62.4             |
| BSK(HI)C6            | 1974    | 3             | 2          | 115.4 ± 10.4 −2.6 ± 7.6 17.7 ± 5.4         | 464.5 ± 141.7   | 49.1            | 7.7               | 54.9             |
| BSK(HI)C7            | 1977    | 3             | 2          | 109.0 ± 12.1 30.6 ± 12.2 17.3 ± 8.4        | 454.0 ± 220.4   | 28.9            | 8.5               | 48.8             |
| Average              |         | 3.3\(^c\)    | 2.1\(^c\)  | 40.0 ± 4.9 8.5 ± 4.3 9.6 ± 3.2             | 250.9 ± 84.3    | 48.8            | 5.0               | 61.9             |

\(^a\) Estimates of components of variance from selection trials in quintals per hectare

\(^b\) Estimates of \(\hat{\sigma}^2_A\) calculated from half-sib progenies and expressed as grams per plant

\(^c\) Values are harmonic means
Table 5.12  Summary of estimates of variance components for grain yield for BSK(S) for eight cycles of S₁ progeny recurrent selection

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Year</th>
<th>No. of trials</th>
<th>Reps/trial</th>
<th>Variance component estimates&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Genetic CV (%)</th>
<th>Average yield (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( \hat{\sigma}^2 )</td>
<td>( \hat{\sigma}^2_{ge} )</td>
<td>( \hat{\sigma}^2_g )</td>
</tr>
<tr>
<td>BSK(S)C0</td>
<td>1954</td>
<td>1</td>
<td>2</td>
<td>13.9 ± 2.2</td>
<td></td>
<td>35.4 ± 6.1</td>
</tr>
<tr>
<td>BSK(S)C1</td>
<td>1958–59</td>
<td>4</td>
<td>2</td>
<td>19.3 ± 1.4</td>
<td>8.6 ± 1.6</td>
<td>38.9 ± 5.9</td>
</tr>
<tr>
<td>BSK(S)C2</td>
<td>1962</td>
<td>3</td>
<td>2</td>
<td>24.5 ± 2.2</td>
<td>10.7 ± 2.7</td>
<td>31.7 ± 5.9</td>
</tr>
<tr>
<td>BSK(S)C3</td>
<td>1965</td>
<td>3</td>
<td>2</td>
<td>22.2 ± 2.1</td>
<td>13.2 ± 2.8</td>
<td>39.5 ± 7.1</td>
</tr>
<tr>
<td>BSK(S)C4</td>
<td>1968</td>
<td>4</td>
<td>2</td>
<td>17.9 ± 1.4</td>
<td>10.6 ± 1.8</td>
<td>18.4 ± 3.3</td>
</tr>
<tr>
<td>BSK(S)C5</td>
<td>1971</td>
<td>4</td>
<td>2</td>
<td>44.8 ± 3.5</td>
<td>13.0 ± 3.4</td>
<td>32.1 ± 5.8</td>
</tr>
<tr>
<td>BSK(S)C6</td>
<td>1974</td>
<td>3</td>
<td>2</td>
<td>65.9 ± 6.0</td>
<td>30.1 ± 7.0</td>
<td>59.3 ± 11.5</td>
</tr>
<tr>
<td>BSK(S)C7</td>
<td>1977</td>
<td>3</td>
<td>2</td>
<td>21.5 ± 2.2</td>
<td>6.6 ± 2.7</td>
<td>38.9 ± 6.8</td>
</tr>
<tr>
<td>Average</td>
<td>2.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2</td>
<td></td>
<td>28.8 ± 2.6</td>
<td>13.2 ± 3.1</td>
<td>30.5 ± 6.6</td>
</tr>
</tbody>
</table>

<sup>a</sup> Estimates of components of variance from selection trials in quintals per hectare

<sup>b</sup> Estimates of \( \hat{\sigma}^2_b \) calculated from S₂ progenies and expressed as grams per plant

<sup>c</sup> Value is harmonic mean
for the other selection programs. It is believed that the following items can partially explain the decrease and then increase of the genetic variance component estimate:

1. Progenies were tested in fewer environments (one or two in most instances) in the first two cycles of selection, and the genotype–environment interaction biased estimates of genetic variance upward.
2. Cycles 2–4 were often tested when environmental conditions were unfavorable because of moisture stress.
3. Before 1970 all test plots were hand harvested with gleaning for dropped ears and ears on broken stalks; since 1970 all plots have been mechanically harvested with no retrieval of any ears not collected by the harvester.

The effect of genotype–environment interaction is found in the summary of recurrent selection programs given in Table 5.13.

After conversion of component of variance estimates from the selection programs, they were averaged to obtain the values at the bottom of Table 5.13. Average estimates of $\hat{\sigma}_{AE}^2$ (166.8 ± 69.2) and $\hat{\sigma}_{A}^2$ (311.2 ± 72.2) show that the estimate of $\hat{\sigma}_{AE}^2$ is 53.6% as great as that of $\hat{\sigma}_{A}^2$. This comparison, therefore, indicates that estimates of $\hat{\sigma}_{A}^2$ obtained in one environment would have on the average a 50% upward bias from genotype–environment interaction. In some instances [e.g., see BS13 and BSSS(R) in Table 5.13] the bias from genotype–environment interaction would be greater than 50% $\hat{\sigma}_{AE}^2$ is 84.2% and 61.8% as large as $\hat{\sigma}_{A}^2$ for BS13 and BSSS(R), respectively. The magnitude of $\hat{\sigma}_{AE}^2$ relative to $\hat{\sigma}_{A}^2$ in Table 5.13 shows that estimates of $\hat{\sigma}_{AE}^2$ are greater for half-sib progenies than for inbred progenies. The five half-sib selection programs (although BS13 includes two cycles of S2 progeny selection) have an average estimate of $\hat{\sigma}_{AE}^2$ that is 66.6% as large as the estimate of $\hat{\sigma}_{A}^2$; whereas the four populations undergoing inbred progeny selection (either S1 or S2) have an average estimate of $\hat{\sigma}_{AE}^2$ that is only 32.8% as large as the average estimate of $\hat{\sigma}_{A}^2$. A more direct comparison of estimates of $\hat{\sigma}_{AE}^2$ relative to $\hat{\sigma}_{A}^2$ can be made for BSK, which is undergoing half-sib and inbred progeny recurrent selection. Although the two selection programs were grown in separate experiments, progenies have been tested in the same years at the same locations each year. Average estimates of $\hat{\sigma}_{A}^2$ are very similar [250.9 for BSK(HI) vs. 241.2 for BSK(S)], but average estimates of $\hat{\sigma}_{AE}^2$ are smaller for BSK(S) [223.0 for BSK(HI) vs. 104.4 for BSK(S)]. Although most average estimates of $\hat{\sigma}_{AE}^2$ are within one standard error of each other, estimates of $\hat{\sigma}_{AE}^2$ tend to be smaller for the inbred progeny selection studies. The evidence shows, however, that genotype–environment interaction is an important factor in estimation of genetic variability. Estimates of $\hat{\sigma}_{g}^2$ obtained in cycles 0 and 1 probably were biased upward.

The reduction in the estimates of $\hat{\sigma}_{g}^2$ for cycles 2–4 across populations were all obtained from data collected during 1954–1963. Moisture stress in evaluation trials was common during this time, which suggests that the stress environment caused a reduction in genetic variability expressed among the half-sib progenies tested but had a small effect on the inbred progenies [BSK(S) in Table 5.12]. Evidence from
<table>
<thead>
<tr>
<th>Selection population</th>
<th>Cycles of selection</th>
<th>Variance component estimates</th>
<th>( \hat{\sigma}^2 )</th>
<th>( \hat{\sigma}_{AE}^2 )</th>
<th>( \hat{h}^2 ) (%)</th>
<th>Genetic CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS13(R)</td>
<td>11</td>
<td>641.4 ± 37.0</td>
<td>156.5 ± 20.1</td>
<td>214.5 ± 64.0</td>
<td>68.1</td>
<td>10.7</td>
</tr>
<tr>
<td>BS15(R)</td>
<td>11</td>
<td>1367 ± 99.7</td>
<td>204.7 ± 107.6</td>
<td>349.3 ± 92.4</td>
<td>52.8</td>
<td>5.4</td>
</tr>
<tr>
<td>BSCB1(R)</td>
<td>11</td>
<td>1560.8 ± 102.3</td>
<td>215 ± 115.4</td>
<td>394.3 ± 93.6</td>
<td>53.8</td>
<td>6.2</td>
</tr>
<tr>
<td>BSCB12(HI)</td>
<td>11</td>
<td>1516.7 ± 107.6</td>
<td>204.7 ± 107.6</td>
<td>349.3 ± 92.4</td>
<td>52.8</td>
<td>5.4</td>
</tr>
<tr>
<td>BSK(HI)</td>
<td>8</td>
<td>1350 ± 157.4</td>
<td>230 ± 112.8</td>
<td>454.2 ± 130.3</td>
<td>54.9</td>
<td>6.2</td>
</tr>
<tr>
<td>BSK(S)</td>
<td>8</td>
<td>1048.9 ± 128.2</td>
<td>223 ± 112.8</td>
<td>454.2 ± 130.3</td>
<td>54.9</td>
<td>6.2</td>
</tr>
<tr>
<td>BS12(HI)</td>
<td>8</td>
<td>1048.9 ± 128.2</td>
<td>223 ± 112.8</td>
<td>454.2 ± 130.3</td>
<td>54.9</td>
<td>6.2</td>
</tr>
<tr>
<td>BS16(S)</td>
<td>5</td>
<td>1048.9 ± 128.2</td>
<td>223 ± 112.8</td>
<td>454.2 ± 130.3</td>
<td>54.9</td>
<td>6.2</td>
</tr>
<tr>
<td>BS10 × BS11(FR)</td>
<td>9</td>
<td>1246.8 ± 87.8</td>
<td>132.4 ± 81.3</td>
<td>444.5 ± 82.6</td>
<td>61.3</td>
<td>9.1</td>
</tr>
<tr>
<td>Average</td>
<td>8.0</td>
<td>926.5 ± 73.3</td>
<td>166.8 ± 69.2</td>
<td>311.2 ± 72.2</td>
<td>66.0</td>
<td>14.9</td>
</tr>
</tbody>
</table>

\* BS13 includes average of estimates for BS13(HT) and BS13(S) (see Table 5.7)
other studies evaluating types of environments for maximizing expression of genetic variability is not extensive (Stevenson, 1965; Arboleda and Compton, 1974). Data from selection trials, however, suggest that genetic variability was reduced in stress environments. Effects of environment seem more plausible than a reduction in variability because of effective selection. In most instances only two or three cycles of selection were completed, and changes in gene frequency were probably not great enough to significantly change the estimates of genetic variance. For BS13(HT), estimates of \( \hat{\sigma}^2_g \) obtained in cycles 3 \((-1.6 \pm 3.6)\) and 4 \((3.4 \pm 1.3)\) in 1955 and 1958–1959, respectively, are significantly smaller than those obtained for cycles 2 and 5 in 1952 and 1962, respectively (Table 5.7). Information in the selection experiments supports the hypothesis that genetic variability is compressed in stress environments. It seems that reduction of genetic variability in the middle cycles of selection was a function of the environments sampled rather than a reduction from changes in gene frequency, particularly for a complex trait such as grain yield. On the other hand, a dramatic increase in estimates of \( \hat{\sigma}^2_g \) in later cycles of selection can be accounted for by the harvesting methods used to collect yield data. Before 1970 all yield test plots were hand harvested, and all ears were included; since 1970, mechanical harvesters have been used to determine grain yields, and, consequently, many ears were not included in yield measurements. Mechanical harvesting, therefore, measures only harvestable yield and does not measure total yield potential of genotypes. The change in harvesting methods for evaluation of progenies in selection studies has opposing views:

1. Mechanical harvesting will not permit the total yield of genotypes unless ears are gleaned from plots by hand.
2. If populations undergoing selection are for use in applied breeding programs, grain yield from standing plants is the important criterion.

The increased genetic variability for grain yield from mechanically harvested cycles of selection probably was caused indirectly by the influence of ear shank attachment and root and stalk lodging, none of which would have influenced relative yields among progenies by use of hand-harvesting procedures.

Genetic coefficients of variation \((\hat{\sigma}_g / \bar{x}) \times 100\) were calculated for each cycle to express the genetic standard deviation as a fraction of the cycle mean. Although the cycle means in Tables 5.7–5.12 are confounded by environmental effects, they exhibit an increasing trend from the original (C0) to the last cycle of selection in all instances. Although comparisons are not valid because of year effects, yields of BS13(HT) increased from 35.1 q/ha for cycle C0 to 67.4 q/ha in cycle C6 (Table 5.7). Seven cycles of half-sib selection nearly doubled yields of half-sib progenies; interestingly, when progeny evaluation was changed from half-sib to S2, average yields of the C0 and C1 cycles of S2 selection were approximately the same as half-sib yield in 1940. Although it is assumed that numerically larger quantities tend to vary more than numerically smaller quantities, the genetic coefficients of variation in Tables 5.7–5.12 do not show a trend with greater yields of later cycles.
of selection. It does not seem that a trend is evident in the size of the genetic coefficients of variation associated with average yields of successive cycles of selection; they are greater, however, when inbred progeny evaluation is used (Table 5.7, last four cycles, and Table 5.12). Genetic coefficients of variation follow the same trend as estimates of $\hat{\sigma}_g^2$ rather than average yields of different cycles.

Although there are instances that suggest some reduction in genetic variance after the first cycles of selection, it seems that the genetic variance estimates in populations undergoing recurrent selection fluctuate among selection cycles (estimates in different environments for each cycle) without evidence of a consistent reducing trend. Consider, for example, estimates of genetic components of variance $\hat{\sigma}_A^2$ and genetic coefficients of variation in Tables 5.8 and 5.9 for BSSS(R) and BSCB1(R) populations that are undergoing reciprocal recurrent selection. After 10 selection cycles estimates of $\hat{\sigma}_g^2$ are similar in cycles C0 and C10 for both populations [BSSS(R): 13.3 ± 3.4 (C0) vs. 14.1 ± 3.6 (C10); and BSCB1(R): 20.8 ± 3.7 (C0) vs. 13.7 ± 3.7 (C10)]. Also, genetic coefficients of variation did not change greatly from C0 to C10, although average yields of C10 testcrosses of BSSS(R) and BSCB1(R) are 59 and 75% greater than C0 populations.

Therefore, data show that selection programs are satisfying the two criteria of successful application of cyclical selection programs for maize improvement:

1. maintenance of genetic variability for future selection and,
2. improvement of overall performance of population crosses in successive cycles of selection.

Estimates of $\hat{\sigma}_A^2$ are given in Tables 5.7–5.12 for each cycle of selection, with adjustments made for half-sib and inbred progenies used in selection programs. Variation of estimates of $\hat{\sigma}_A^2$ among cycles of selection follows the same trend as estimates of components of genetic variance $\hat{\sigma}_g^2$. Because there is no evidence that additive genetic variance has changed among cycles of selection, estimates of $\hat{\sigma}_A^2$ averaged for each population are summarized in Table 5.13. Additionally, estimates for three populations that have had only two and three cycles of selection are included for comparison. BS13(HT) and BSSS(R) involve the same population but two different recurrent selection procedures, and estimates of $\hat{\sigma}_A^2$ in Table 5.13 are very similar to those in Table 5.5 for BSSS. Estimates of $\hat{\sigma}_A^2$ obtained from selection experiments (Table 5.13) are within one standard error of the average of estimates obtained from mating designs in Table 5.5. Hence it seems that valid estimates of $\hat{\sigma}_A^2$ can be obtained from long-term selection experiments, provided sufficient sampling, testing, and cycles of selection are available. Experimental errors are inherently large for variance component estimates, and estimates obtained from a particular cycle of selection may deviate considerably from the average [see cycle C3 for BS13(HT) in Table 5.7 and cycle C4 for BSSS(R) in Table 5.8]. Sampling errors and bias from genotype–environment interaction could cause erroneous conclusions from only one cycle. Only when sufficient cycles of selection, sampling, and testing have been completed will valid estimates be obtained. Usually, only 100 progenies were tested in each cycle of selection for BS13 and BSSS(R), and errors
of estimation were about twice those obtained from mating designs that included 231 to 480 progenies for testing. Estimates of heritability, on a progeny mean basis, from the selection studies for BS13(HT) (40.7%) and BSSS(R) (52.8%) also were similar to average heritability of the four estimates included in Table 5.5 (41.4%).

The BSSS [BS13 and BSSS(R)] and BSK [BSK(HI) and BSK(S)] populations seem to have less genetic variability than other populations (Table 5.2). For the two selection studies involving BSSS and BSK, average estimates of $\hat{\sigma}_A^2$ are 286.9 and 246.0, respectively; whereas the average estimate of $\hat{\sigma}_A^2$ for the other six populations is 341.1. The two estimates for BSSS, however, are in good agreement with the average of all estimates of $\hat{\sigma}_A^2$ for synthetic populations in Table 5.2. The two estimates of BSK [250.9 for BSK(HI) and 241.2 for BSK(S)] also are in agreement with that ($\hat{\sigma}_A^2 = 202$) reported by Wright et al. (1971) by use of weighted least squares analysis of progenies developed from the diallel and triallel mating designs. Both BSSS and BSK are synthetic varieties developed by recombination of 16 and 8 lines, respectively. The 16 lines used to form BSSS were characterized as having strong stalks and thus originated from several sources. The 8 lines used to form BSK, however, originated from the Krug open-pollinated variety and were selected on the basis of testcross yields with the parental variety Krug used as the tester (Lonquist, 1949). Genetic variability in BSSS and BSK may have been limited by either the restricted sample of lines included in their synthesis or a greater gene frequency. BS12 is an open-pollinated variety of unknown origin that has the largest genetic variability (about twice), based on average estimates from selection studies (Table 5.13). BS16, BSTL, and BS2 have had fewer cycles of selection completed, and estimates of $\hat{\sigma}_A^2$ were similar to those for BSSS and BSK. BSTL was developed from a variety cross, Lancaster X Tuxpeno (a Mexican variety), and backcrossed to Lancaster; its estimate of $\hat{\sigma}_A^2$ was similar to that for BSSS. BS16 has the smallest estimate of $\hat{\sigma}_A^2$. BS16 was developed by six cycles of mass selection for early flowering in ETO Composite, obtained from Colombia, South America (Hallauer and Sears, 1972). Limited evidence from BS16 suggests that selection within exotic germplasm reduced genetic variability available to the breeder.

Comparison of estimates of $\hat{\sigma}_A^2$ from selection experiments with those obtained from use of mating designs shows that they are equally valid. If estimates of $\hat{\sigma}_A^2$ from selection experiments were included in Table 5.2 and Fig. 5.1, they would not deviate from estimates obtained from use of mating designs. Figure 5.1 shows that all types of populations are expected to have sufficient additive genetic variance to show response from selection, though some more than others. Estimates from selection experiments show variation among cycles of selection, but the range of estimates is not any greater than those obtained from individual studies by use of mating designs. Sampling and genotype–environment interaction probably account for variation of estimates among cycles of selection; these factors, however, are just as important if insufficient sampling and testing are used to obtain estimates from mating designs. Often, selection experiments had smaller sample sizes but not always, especially if progenies across cycles are considered but testing is more extensive during selection studies. However, summaries of estimates given in Tables 5.1, 5.2, and 5.13 were obtained from data collected during the same period. Selection experiments were
conducted cyclically since the 1940s and use of mating designs to estimate genetic parameters were extensive after the publication of Comstock and Robinson (1948). Again, summaries of both methods show similar results. Use of mating designs will provide estimates of genetic parameters more quickly if sufficient sampling and testing are used. Silva and Hallauer (1975) obtained the estimate of $\hat{\sigma}_A^2 = 166 \pm 24$ for yield in BSSS in 4 years (experiment 3 in Table 5.5); on the other hand, the estimate for yield of $\hat{\sigma}_A^2 = 214.5 \pm 64.0$ was obtained from 11 cycles of reciprocal recurrent selection in BSSS, which required 27 years. Nowadays, with winter nurseries allowing three generations per season, this same program could be conducted in < 10 years (e.g., still two times quicker with mating designs). Although the estimates are similar, they are not exactly comparable because Silva and Hallauer (1975) used intra-population progenies, whereas the BSCB1 population was used as tester for the BSSS population in the reciprocal recurrent selection program. About 800 progenies were evaluated in both instances; 800 full-sib progenies were tested that were developed from designs I and II mating designs and about 800 half-sib progenies (8 cycles × 100 half-sib progenies per cycle) were tested from the selection experiment. Estimates of $\hat{\sigma}_A^2$ from the selection experiment were obtained, however, in 23 environments vs. only 6 environments for the mating designs. However, if we use standard errors of estimates of $\hat{\sigma}_{AE}^2$ and $\hat{\sigma}_A^2$ to determine significance, estimates of $\hat{\sigma}_{AE}^2$ and $\hat{\sigma}_A^2$ are not significantly different by the two methods of estimation.

Estimates of genetic parameters in maize can be obtained more quickly from use of mating designs than from selection experiments. If the experimenter wants to determine the genetic variability of a population to predict future progress from selection, the use of adequate mating designs becomes necessary. If the experimenter is patient and is willing to collect data from several cycles of selection before predicting progress from selection it is not necessary to develop progenies from use of mating designs and conduct special studies to obtain this information, which can be quite costly if adequate sampling and testing are included. To answer specific questions about a particular population for its potential use in a breeding program, it is also necessary to use some type of mating design. Currently, there are not many applied breeding programs conducting long-term selection programs to maximize genetic improvement of germplasm adapted to certain regions. Other priorities have been emphasized in both the private and the public sectors although scientists still have the option to decide. In addition to providing unique, adapted, and improve germplasm to marginal and non-marginal environments, an important reason to keep conducting them is that data from different long-term recurrent selection programs include the evaluation of different types of progeny. Based on data available from the combined analyses of variance, estimates of the genetic parameters in maize can be obtained.

### 5.4 Epistasis Variance and Effects

Epistasis should be present in exceptional hybrids derived from a specific combination of elite lines. Epistatic interactions seem to be specific and, therefore,
not extensively present across hybrids. Results from epistasis genetic variance studies, however, have challenged breeders to detect it, especially when measured in populations.

5.4 Epistasis Variance and Effects

Estimates of additive genetic variance and variance due to dominance deviations for maize populations (Table 5.1) were obtained under the assumption of no epistasis. In most instances only one or two equations were available for estimation of genetic components of variance. Estimation, therefore, frequently was limited by the mating design used:

1. If only one source of variation was available (e.g., variability among full-sib families), all non-additive sources (both dominance and epistasis) of variation were assumed absent.
2. If two sources of variation were available (e.g., variability among half-sibs and full-sibs within half-sibs), epistasis was assumed absent.

The assumptions were necessary because of limitations of mating designs used to develop progenies for test. The simpler the mating design, the greater the restrictions needed for estimation. Fortunately, in maize it does not seem that restrictions imposed for the assumption of no epistasis seriously biased estimates of additive genetic and dominance variance components.

Development of more complex mating designs permitted estimation of additional components of genetic variance. But adding epistasis to a model must increase predictive power (Dudley and Johnson, 2009). In most instances the more complex mating designs permitted estimation of all types of digenic epistasis and, in a few instances, trigenic epistasis, e.g., additive by additive by additive epistasis. The primary objective of the more complex mating designs was to develop additional covariances of relatives to permit estimation of additional components of genetic variation. One of the first suggestions for estimation of epistatic variances was by Cockerham (1956); designs I and II mating designs were used with parents at two different levels of inbreeding, but the progenies evaluated were non-inbred in both instances. The procedure suggested was used by Eberhart (1961) and Silva (1974), Rawlings and Cockerham (1962a, b) developed the triallel and quadrallel analyses that provide up to nine covariances of relatives; these analyses permitted $F$-tests for the presence of epistasis in the analyses of variance and estimation of epistatic components of variance. Wright (1966) used diallel and triallel analyses, which provided nine mean squares, for estimation of epistasis in Krug Hi I Synthetic 3. Chi (1965) used a complex mating design suggested by Kempthorne (1957, pp. 425–26) that included 11 variances and 55 covariances among relatives to estimate epistasis in an open-pollinated variety Reid Yellow Dent. Epistasis was recently estimated utilizing the triple testcross mating design (Melchinger et al., 2008), recombinant inbred lines (Buckler et al., 2009), and improved populations (Dudley, 2008) across economically important quantitative traits.

Estimation of epistatic components of variance has not been generally satisfying. Most of the studies included adequate sampling and testing, but the results of estimation have been disappointing but seem promising. Epistasis for a complex trait,
Hereditary Variance: Experimental Estimates

such as yield, is expected to exist and additive types of epistasis would be useful to the breeder in selection programs. However, realistic estimates of additive by additive epistasis have not been obtainable. Hence, genetic models used could be inadequate and epistatic variance could be small relative to total genetic variance of maize populations. As shown in Chapter 2, it seems that a major problem is the inherent correlation of coefficients of epistatic components of variance with those of additive genetic and dominance variance components (see Chapter 4).

Estimates of epistatic variance for five maize populations from four independent studies are included in Table 5.14. Three populations (Jarvis, Indian Chief, and Reid Yellow Dent) are open-pollinated varieties, whereas the other two [BSK (a strain of Krug Yellow Dent) and BSSS] are synthetic varieties. Each of the studies used different methods, but they all concluded that realistic estimates of epistatic components of variance were not obtainable and they all reverted to simpler genetic models for estimation of additive and dominance components of genetic variance. In some instances, negative estimates of epistatic variance components were two times greater than their standard errors. However, changing probability levels for inclusion of epistatic effects in genetic models was recently proposed by Dudley and Johnson (2009).

Eighteen estimates of epistatic variance components are included in Table 5.14, and 11 of the estimates are negative. None of the positive estimates is greater than two times its standard error and most are within one standard error. Eberhart et al. (1966) concluded, ‘Additive variance appeared to account for the largest proportion of the total genetic variance for all characters in both varieties.’ Chi et al. (1969), from the complex mating design used in Reid Yellow Dent, concluded, ‘The results indicated that epistatic variances were negligible in relation to the additive and dominance variance components for the seven characters studied. The high correlations among the coefficients of the genetic parameters inevitably reduced the sensitivity for detecting epistasis.’ Wright et al. (1971) used the maximum likelihood estimation method on the mean squares from diallel and triallel analyses to fit the error and a six-parameter genetic model; and they concluded, ‘It was not possible to obtain realistic estimates of the epistatic components of variance, although significant effects were detected in the analysis of variance. For the two-parameter genetic model, the largest proportion of the total genetic variance was additive for all traits.’ Silva and Hallauer (1975), from an extensive study of BSSS, concluded, ‘Epistatic variance was not an important component of the genetic variance for yield in BSSS. Additive genetic variance accounted for 93.2% of the total variability. Including the variance due to dominance deviations in the models accounted for more than 99% of the variation, with no improvement when additive by additive epistatic variance was included. Estimates of epistatic variances obtained from models that included more than one digenic epistatic component usually were negative and unrealistic with much greater standard errors. The use of a complete model (as many epistatic terms as permitted by the number of independent equations) made the X-matrix nearly singular.’ As a consequence, each study resorted to the two-parameter genetic model for estimation of additive genetic and dominance components of variance. Estimates and their interactions with environments for the simpler models are shown
Table 5.14  Estimates of epistatic components of variance for yield by use of different methods of estimation for five maize populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Epistatic variance components</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{\sigma}^2_{Ig})</td>
<td>(\hat{\sigma}^2_{Ia})</td>
<td>(\hat{\sigma}^2_{Id})</td>
<td>(\hat{\sigma}^2_A)</td>
<td>(\hat{\sigma}^2_D)</td>
</tr>
<tr>
<td>Jarvis (Eberhart et al., 1966)(^a)</td>
<td>(-300 \pm 136)</td>
<td>(128 \pm 119)</td>
<td>(-428 \pm 187)</td>
<td>(640 \pm 113)</td>
<td>(407 \pm 156)</td>
</tr>
<tr>
<td>Indian Chief (Eberhart et al., 1966)(^\text{b})</td>
<td>(175 \pm 132)</td>
<td>(78 \pm 107)</td>
<td>(97 \pm 183)</td>
<td>(313 \pm 80)</td>
<td>(247 \pm 138)</td>
</tr>
<tr>
<td></td>
<td>(\hat{\sigma}^2_{AA})</td>
<td>(\hat{\sigma}^2_{AD})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reid Yellow Dent (Chi et al., 1969)(^\text{b})</td>
<td>8.12</td>
<td>5.40</td>
<td>7.04</td>
<td>646.70</td>
<td>37.73</td>
</tr>
<tr>
<td></td>
<td>(\hat{\sigma}^2_{AA})</td>
<td>(\hat{\sigma}^2_{AD})</td>
<td>(\hat{\sigma}^2_{DD})</td>
<td>(\hat{\sigma}^2_{AAA})</td>
<td></td>
</tr>
<tr>
<td>BSK (Wright et al., 1971)(^\text{c})</td>
<td>(-40 \pm 225)</td>
<td>182 (\pm 236)</td>
<td>(-225 \pm 96)</td>
<td>(-164 \pm 257)</td>
<td>221 (\pm 54)</td>
</tr>
<tr>
<td>Unweighted</td>
<td>(132 \pm 286)</td>
<td>291 (\pm 270)</td>
<td>(-102 \pm 120)</td>
<td>(-100 \pm 335)</td>
<td>265 (\pm 40)</td>
</tr>
<tr>
<td>Weighted</td>
<td>(-94 \pm 165)</td>
<td>(-305 \pm 122)</td>
<td></td>
<td></td>
<td>271 (\pm 73)</td>
</tr>
<tr>
<td>BSSS (Silva, 1974)(^\text{d})</td>
<td>(-94 \pm 165)</td>
<td></td>
<td>(-204 \pm 81)</td>
<td></td>
<td>271 (\pm 73)</td>
</tr>
</tbody>
</table>

\(^a\) Estimates obtained by mathematical arrangements of genetic expectations of the components of variance, where \(\hat{\sigma}^2_{Ig}\) includes all epistatic components of variance, \(\hat{\sigma}^2_{Ia}\) includes only additive types of epistatic variance, and \(\hat{\sigma}^2_{Id}\) includes all types of epistatic variance except additive by additive.

\(^b\) Analysis of variance estimates of genetic components of variance averaged over 2 years \((\times 10^{-2})\).

\(^c\) Estimates obtained by ordinary least squares and weighted least squares methods of estimation.

\(^d\) Estimates obtained by the maximum likelihood method of estimation.
in Table 5.15. In most instances estimates of genetic parameters in Table 5.15 are smaller and have smaller standard errors than those obtained from genetic models that included epistasis (Table 5.14). All estimates of additive genetic variance and their interactions with environments exceeded twice their standard errors. In most instances estimates of dominance also exceeded twice their standard errors. If we consider estimates that exceed twice their standard errors as being significantly different from zero, the experiments seemed to have used adequate sampling and testing procedures for estimation of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) genetic parameters. Moreover, estimates of \( \hat{\sigma}_A^2 \) in Table 5.15 are within sampling error of those given in Tables 5.1 and 5.2.

As an example, below is the estimation of epistatic components of genetic variance from the combined analysis on BSSS. Mating designs I and II were utilized to develop 800 full-sib progenies (480 design I and 320 design II full-sib progenies) that were evaluated in six environments. From the genetic expectations of mean squares for the designs I and II combined analysis of variance, the following estimates were calculated.

### Design I \((F = 0)\)
- \( \hat{\sigma}_m^2 = \text{Cov} \ HS = 34.8 \pm 11.4 \)
- \( 4\hat{\sigma}_m^2 = \hat{\sigma}_A^2 = 139.3 \pm 45.5 \)
- \( \hat{\sigma}_f^2 = \text{Cov} \ FS - \text{Cov} \ HS = 101.9 \pm 10.0 \)
- \( 4(\hat{\sigma}_f^2 - \hat{\sigma}_m^2) = \hat{\sigma}_D^2 = 268.2 \pm 60.5 \)
- \( \hat{\sigma}_w^2 = \hat{\sigma}_{we}^2 + (\hat{\sigma}_G^2 - \text{Cov} \ FS) = 1654.1 \pm 23.3 \)
- \( \hat{\sigma}_p^2 = 354.3 \pm 9.5 \)

### Design II \((F = 1)\)
- \( \hat{\sigma}_m^2 = \text{Cov} \ HS = 37.1 \pm 18.0 \)
- \( 2\hat{\sigma}_m^2 = \hat{\sigma}_A^2 = 74.2 \pm 35.9 \)
- \( \hat{\sigma}_f^2 = \text{Cov} \ HS = 84.5 \pm 26.3 \)
- \( \hat{\sigma}_f^2 = \text{Cov} \ FS - \text{Cov} \ HS_f - \text{Cov} \ HS_m = 167.2 \pm 21.8 \)
- \( \hat{\sigma}_w^2 = \hat{\sigma}_{we}^2 = 1280.6 \pm 29.6 \)
- \( \hat{\sigma}_p^2 = 332.3 \pm 11.0 \)

Additional calculations include estimates of \( \hat{\sigma}_{wg}^2 \) from design I and \( \hat{\sigma}_{HS}^2 \) from design II. Because \( \hat{\sigma}_{wg}^2 \) (in design II) includes only the full-sib progeny measurement and plant-to-plant environmental errors, an estimate of the within-plot genetic variance can be obtained as 1654.1 (from design I)–1280.6 (from design II), which becomes 373.5 ± 32.1, the estimate of \( \hat{\sigma}_{wg}^2 \) or \( \hat{\sigma}_G^2 - \text{Cov} \ FS \). The value of \( \hat{\sigma}_{HS}^2 \) was calculated from the Design II analysis by pooling degrees of freedom and sums of squares for male and female sources of variation; \( \hat{\sigma}_{HS}^2 \) was obtained from expected mean squares as 60.8 ± 16.2, which is the average of \( \hat{\sigma}_m^2 \) and \( \hat{\sigma}_f^2 \) estimates. Hence, an estimate of \( \hat{\sigma}_A^2 \) from design II is 121.6 ± 32.4.

If we assume no epistasis, the two independent estimates of \( \hat{\sigma}_A^2 \) are 139.3 ± 45.5 (design I) and 121.6 ± 32.4 (design II) and the estimates of \( \hat{\sigma}_D^2 \) are 268.2 ± 60.5 (design I) and 167.2 ± 21.8 (design II). The estimates of \( \hat{\sigma}_A^2 \) and \( \hat{\sigma}_D^2 \) are within two standard errors of each other but the estimates from design I have greater standard errors, particularly the estimate of \( \hat{\sigma}_D^2 \). Before pooling male and female half-sibs
Table 5.15 Estimates of additive and dominance genetic variances for yield with the assumption of no epistasis for five maize populations

<table>
<thead>
<tr>
<th>Populations</th>
<th>( \hat{\sigma}^2_A )</th>
<th>( \hat{\sigma}^2_{AE} )</th>
<th>( \hat{\sigma}^2_D )</th>
<th>( \hat{\sigma}^2_{DE} )</th>
<th>( \hat{\sigma}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis, original Eberhart et al. (1966)</td>
<td>247 ± 105</td>
<td>234 ± 95</td>
<td>574 ± 183</td>
<td>202 ± 193</td>
<td>1045 ± 35</td>
</tr>
<tr>
<td>Jarvis, reconstituted Eberhart et al. (1966)</td>
<td>374 ± 58</td>
<td>214 ± 45</td>
<td>127 ± 33</td>
<td>150 ± 43</td>
<td>1018 ± 33</td>
</tr>
<tr>
<td>Indian Chief, original Eberhart et al. (1966)</td>
<td>181 ± 97</td>
<td>134 ± 105</td>
<td>43 ± 179</td>
<td>407 ± 138</td>
<td>1108 ± 41</td>
</tr>
<tr>
<td>Indian Chief, reconstituted Eberhart et al. (1966)</td>
<td>259 ± 45</td>
<td>189 ± 39</td>
<td>140 ± 31</td>
<td>25 ± 41</td>
<td>1199 ± 37</td>
</tr>
<tr>
<td>Reid Yellow Dent Chi et al. (1969)</td>
<td>239 ± 41</td>
<td>—</td>
<td>329 ± 102</td>
<td>—</td>
<td>339 ± 8</td>
</tr>
<tr>
<td>BSK, unweighted Wright et al. (1971)</td>
<td>181 ± 16</td>
<td>102 ± 27</td>
<td>85 ± 47</td>
<td>-29 ± 90</td>
<td>538 ± 78</td>
</tr>
<tr>
<td>BSK, weighted Wright et al. (1971)</td>
<td>202 ± 88</td>
<td>105 ± 38</td>
<td>67 ± 35</td>
<td>56 ± 37</td>
<td>393 ± 21</td>
</tr>
<tr>
<td>BSSS Silva and Hallauer (1975)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML(^a)</td>
<td>169 ± 24</td>
<td>92 ± 10</td>
<td>193 ± 21</td>
<td>75 ± 12</td>
<td>185 ± 7</td>
</tr>
<tr>
<td>LS</td>
<td>138 ± 34</td>
<td>149 ± 63</td>
<td>73 ± 147</td>
<td>-6 ± 303</td>
<td>355 ± 324</td>
</tr>
<tr>
<td>WLS</td>
<td>150 ± 34</td>
<td>91 ± 16</td>
<td>189 ± 30</td>
<td>76 ± 17</td>
<td>185 ± 11</td>
</tr>
</tbody>
</table>

\(^a\) ML, LS, and WLS refer to the maximum likelihood, least squares, and weighted least squares methods of estimation, respectively
in design II, the standard errors of the estimates of $\hat{\sigma}^2_A$ were similar for both mating designs. Because the coefficient of inbreeding is different for the parents of design I ($F = 0$) and design II ($F = 1$), the covariances of half-sibs have different genetic expectations. If we include additive by additive epistasis $\hat{\sigma}^2_{AA}$ in the genetic expectations, $\text{Cov HS}$ for design I is $(\gamma_4)\hat{\sigma}^2_A + (\gamma_{10})\hat{\sigma}^2_{AA}$ and for design II is $(\gamma_2)\hat{\sigma}^2_A + (\gamma_4)\hat{\sigma}^2_{AA}$. Coefficients of genetic expectations of the covariances of half-sibs are different; hence we can use expectations and observed values of the two covariances of half-sibs to estimate $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_{AA}$:

$$\hat{\sigma}^2_{HS} = (\gamma_2)\hat{\sigma}^2_A + (\gamma_4)\hat{\sigma}^2_{AA} = 60.8 \quad \text{(design II)}$$

and,

$$\hat{\sigma}^2_m = (\gamma_4)\hat{\sigma}^2_A + (\gamma_{10})\hat{\sigma}^2_{AA} = 34.8 \quad \text{(design I)}$$

By solving the two equations, we find the estimate of $\hat{\sigma}^2_A$ is 156.8 and the estimate of $\hat{\sigma}^2_{AA}$ is $-70.8$. The negative estimate of $\hat{\sigma}^2_{AA}$, however, is not significant because the large standard error (223.54) indicates that it is not different from zero. This estimate of $\hat{\sigma}^2_{AA}$ was obtained by use of two estimates of $\text{Cov HS}$ from different mating designs and shows no evidence of epistatic variance. Another estimate of $\hat{\sigma}^2_A$ can be calculated from the three sources of genetic variation in the design I analysis. Because within-plot variation was available from both mating designs, an estimate of the within-plot genetic variation can be obtained; $\hat{\sigma}^2_{WG}$ is 373.5 ± 32.1, which is the total genetic variance minus $\text{Cov FS}$. If we include $\hat{\sigma}^2_{AA}$ in the expectations for the design I analysis, we have

$$\hat{\sigma}^2_m = (\gamma_4)\hat{\sigma}^2_A + (\gamma_{10})\hat{\sigma}^2_{AA} = 34.8$$

$$\hat{\sigma}^2_{lm} = (\gamma_4)\hat{\sigma}^2_A + (\gamma_4)\hat{\sigma}^2_D + (\gamma_{10})\hat{\sigma}^2_{AA} = 101.9$$

$$\hat{\sigma}^2_{wg} = (\gamma_2)\hat{\sigma}^2_A + (\gamma_4)\hat{\sigma}^2_D + (\gamma_4)\hat{\sigma}^2_{AA} = 373.5$$

The solution of these three equations gives the following estimates of the three components of genetic variance:

$$\hat{\sigma}^2_A = 36.6, \quad \hat{\sigma}^2_D = 63.2 \quad \text{and}, \quad \hat{\sigma}^2_{AA} = 410.4$$

The value of $\hat{\sigma}^2_{AA}$ is very large, while $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$ are smaller than any of the previous estimates. By use of the same three equations, $\hat{\sigma}^2_{AA}$ was deleted and $\hat{\sigma}^2_{DD}$ included in the genetic expectations; the estimates were determined to be $\hat{\sigma}^2_A = 139.2$, $\hat{\sigma}^2_D = 234.2$, and $\hat{\sigma}^2_{DD} = 136.8$. The inclusion of dominance by dominance epistasis increased the estimates of $\hat{\sigma}^2_A$ and $\hat{\sigma}^2_D$ to the levels at which epistasis was assumed
5.4 Epistasis Variance and Effects

absent. Additive by additive epistasis seemed to have a greater effect on $\hat{\sigma}_A^2$ and $\hat{\sigma}_D^2$ variance components than dominance by dominance epistasis for equations from the design I analysis. Both estimates of $\hat{\sigma}_A^2$ given above are biased by epistatic variances, but it seems that the bias is greater when $\hat{\sigma}_{AA}^2$ is not included in the model. The estimate of $\hat{\sigma}_{AA}^2$ from use of covariances of half-sibs from designs I and II analyses, however, showed no evidence of additive by additive epistasis.

Discussion of estimates of epistatic variance has been limited to estimates reported for yield. Estimates from classical quantitative genetic studies reported for different plant and ear traits, however, show similar results. Molecular marker studies detected epistatic interactions for days to male and female flowering only within two families and for anthesis–silk interval which explained only 2% of the total phenotypic variation (Buckler et al., 2009). Dudley and Johnson (2009) followed a different approach in which the importance of epistasis for grain oil and protein suggests that interacting gene networks may be important though reports of epistatic interactions from marker data are limited. Moreover, even though quantitative estimates of epistatic variance in maize populations have not been convincing, reports have indicated that epistatic effects (e.g., based on means and not variances) are present in quantitative traits. Most of the evidence was obtained by use of mean comparisons, which included observed and predicted performance of different types of hybrids (single, three-way, and double-cross), generations having different levels of percent homozygosity, and comparisons of different generations of inbreeding with theoretical levels of homozygosity. In most instances the generations were derived by crossing two inbred lines and one would expect certain hybrid combinations will have specific desirable epistatic effects when compared to populations. Genetic models proposed by Anderson and Kempthorne (1954), Cockerham (1954), Hayman and Mather (1955), and Hayman (1958, 1960) permit estimation of additive, dominance, and epistatic gene effects that are based on the factorial model used in the design of experiments. In all instances, qualitative evidence rather than quantitative evidence of epistasis is available.

Gamble (1962a, b) obtained estimates of six genetic parameters ($m$, $a$, $d$, $aa$, $ad$, and $dd$) from six generations ($P_1$, $P_2$, $F_1$, $F_2$, $BC_1$ from the cross $P_1 \times F_1$, and $BC_2$ from the cross $P_2 \times F_1$) generated from crosses among six inbred lines of maize. All lines had good general combining ability. Estimates of the six genetic parameters for the 15 crosses obtained from generation means of four experiments were obtained for six traits. Frequency of significant effects is summarized in Table 5.16 for each trait. Dominance effects were significant in all crosses for all traits except kernel-row number. Occurrence of significant additive effects also was high for all except yield, which was significant in 47% of the 15 crosses. It seems that additive and dominance effects made a significant contribution to inheritance of these traits for this particular set of crosses. Although not as frequent as additive and dominance effects, significant epistatic effects were frequent for all traits. Plant height had the greatest frequency of significant digenic epistatic effects among traits and yield the least. Of the three digenic epistatic effects, $a \times d$ had the greatest occurrence and $a \times a$ the least. Gamble concluded that all gene effects contributed to the inheritance
Table 5.16  Number of significant (0.05 and 0.01 probability levels) estimates of six genetic effects for 15 crosses among six inbred lines (Gamble 1962a, b). Relative frequency of the total significant effects for all traits shown in last column

<table>
<thead>
<tr>
<th>Genetic effect</th>
<th>Yield</th>
<th>Kernel-row number</th>
<th>Ear length</th>
<th>Ear diameter</th>
<th>Seed weight</th>
<th>Plant height</th>
<th>Relative frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>Additive (a)</td>
<td>7</td>
<td>12</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>73</td>
</tr>
<tr>
<td>Dominance (d)</td>
<td>15</td>
<td>9</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>93</td>
</tr>
<tr>
<td>a × a</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>a × d</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>48</td>
</tr>
<tr>
<td>d × d</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>33</td>
</tr>
</tbody>
</table>

of traits studied. For yield, estimates of dominance gene effects were quite important and additive effects were low in magnitude and often non-significant. All dominance effect estimates for yield were positive, whereas nine of the $d \times d$ estimates were negative. All except three of the $a \times a$ estimates were positive. Darrah and Hallauer (1972) and Sprague and Suwantaradon (1975) found similar results for yield and other traits. Hayman (1960), however, pointed out that the presence of epistatic effects would bias estimates of additive and dominance effects.

Moll et al. (1963) used generation mean analysis to study inheritance of resistance to brown spot (Physoderma maydis) in six crosses among four inbred lines. Significant epistatic effects were detected in two crosses, additive effects in three crosses, and non-significant dominance effects in all crosses. Because of the genetic structure of populations studied, estimates of additive and dominance genetic variances also were obtained. The estimates of $\hat{\sigma}^2_A$ exceeded twice their standard errors in four of the six crosses, but no estimates of $\hat{\sigma}^2_D$ exceeded twice their standard errors. Moll et al. (1963) concluded that much genetic variation in brown spot reaction is additive, but heritability seems low. Hughes and Hooker (1971) reported similar results in a study of four crosses for the nature of gene action conditioning resistance to northern leaf blight (Helminthosporium turcicum Pass.). Significant epistatic effects were noted in all crosses by generation mean analysis, but additive genetic effects were of major importance in three of the four crosses. Assuming no epistasis, estimates of $\hat{\sigma}^2_A$ were relatively large and significant, whereas estimates of $\hat{\sigma}^2_D$ were close to zero. They concluded that northern leaf blight resistance was conditioned by a relatively low number of genes with primarily additive effects. Hallauer and Russell (1962) also included genetic populations developed from a cross of two inbred lines that permitted estimation of genetic variances and effects (e.g., both variances and means) for days from planting to flowering and grain moisture and seed weight at physiological maturity. Significant epistatic effects were noted in six of the nine instances with estimates of dominance effects greater than additive effects. Assuming no epistasis, estimates of $\hat{\sigma}^2_D$ exceeded their standard errors but
estimates of $\hat{\sigma}_A^2$ were zero for seed weight and grain moisture at physiological maturity. The estimate of $\hat{\sigma}_A^2$ for days to flowering was significant but the estimate of $\hat{\sigma}_D^2$ was zero (Hallauer, 1965); this is in contrast to the generation mean analysis that showed a positive and significant estimate of dominance effects and a small non-significant estimate of additive effects. As there is little relation in the magnitude of the two sets of estimates for this single cross the knowledge on both variance and means may allow us to focus on other crosses to expect better estimates for selection in the long term. Again, means are estimating the sum of the unsigned genetic effects and the variances are the square of these genetic effects. In addition, even though the analysis of means is more accurate, the fact that epistatic effects were present can make an opposite effect to the precision of the model.

Epistatic effects can be detected by examining the relationship between level of heterozygosity and performance of the quantitative trait. The basis for this relation is based upon Wright’s (1922) studies of hybrid vigor and inbreeding depression: If the change in performance is proportional to the change in heterozygosity, epistasis is either negligible or non-detectable and implies that the change is dependent on some level of dominance. This relation has been examined in maize by Kiesselbach (1933), Neal (1935), Stringfield (1950), Sentz et al. (1954), Robinson and Cockerham (1961), and Martin and Hallauer (1976); the first four reports generally supported the linear relation, whereas the last two reports found instances of a curvilinear (or epistasis) relation between percent heterozygosity and performance. A linear response does not mean an absence of epistatic effects, but it shows no net epistatic effects because cancellation of positive and negative epistatic effects could produce a linear relation. Although some conflicting results have been reported from use of the heterozygosity–performance relation to obtain evidence for the presence of epistasis, it seems reasonable that this should be expected because of the sample of parents included in the studies, environments used for testing, and traits measured. For example, Martin and Hallauer (1976) studied 21 crosses in each of four groups (first cycle, second cycle, good lines, and poor lines) of inbred lines. Each group included seven inbred lines and all possible F1 crosses were made among the seven inbred lines. For each of the 21 crosses within each group, F2, F3, backcross, and backcross-selfed generations were included to form four levels of heterozygosity (0, 25, 50, and 100%). Epistasis was detected most frequently for ear diameter and kernel-row number and least frequently for grain yield for the four groups of lines, with poor lines having greatest frequency of epistasis for all traits. Only 6 instances out of 84 possible were significant at the 0.05% probability level for yield across all groups, which would be expected to occur by chance. Figure 5.2 shows a linear relation of yield with percent heterozygosity, whereas ear diameter has a slightly curvilinear relation. Linear sums of squares accounted for 98.2% and 92.8% of total variation for yield and ear diameter, respectively. Cumulative totals for occurrences of significant epistasis for the four groups were six for grain yield, while 43 for ear diameter. Results of Martin and Hallauer (1976) also agreed with those of Sentz et al. (1954) for effects of environments. They found that combining data across environments made the response to added heterozygosity more linear and decreased the number of crosses exhibiting significant epistatic effects.
Fig. 5.2 Relation of grain yield and ear diameter with level of heterozygosity generations produced from four types of inbred lines (Martin and Hallauer, 1976)

A curvilinear response to increased homozygosity by inbreeding also is evidence for the presence of epistasis, where inbreeding is compared for different generations of unselected lines developed by single-seed descent (Hallauer and Sears, 1973; Cornelius and Dudley, 1974). Linear regressions were highly significant in all instances, indicating additive gene action with some level of dominance. Hallauer and Sears (1973) measured 10 traits and linear regression accounted for 92.9% for ear-row number to over 99% of total variation for plant and ear height and grain yield. Hence the linear regression model accounted for most of the variation among generations, with less than 2% of total variation accounted for by quadratic and deviation sums of squares for 6 of the 10 traits. This also indicates that epistatic effects were of minor importance. Sing et al. (1967) also used unselected lines developed from Jarvis and Indian Chief open-pollinated varieties to develop hybrid progenies representing seven levels of inbreeding. Linear mean squares were significant in all instances. For yield, for example, linear sums of squares accounted for 95.6% and 96.0% of total variation for Indian Chief and Jarvis, respectively. Sing et al. (1967) concluded that ‘if epistatic effects are contributing to the variability in these two populations they must be of the type which leads to a linear relationship between average performance and inbreeding.’

Estimates of relative importance of epistatic effects have been obtained by comparing observed and predicted means of single, three-way, and double-cross hybrids produced usually from elite inbred lines. Unique epistatic complexes of genes may be important in single-cross performance that are specific for each hybrid, particularly if elite inbred lines are used. Production of three-way and double-cross hybrids affords an opportunity for genetic recombination in parental single crosses used to produce the more complex hybrids. Recombination in parental single crosses would disrupt unique gene complexes and cause differences between predicted and observed performances. Comparisons of three-way and double-cross performances, therefore, would provide a qualitative assessment of relative importance of epistatic effects of parental single-cross performance. Bauman (1959), Gorsline (1961), Sprague et al. (1962), Sprague and Thomas (1967), Eberhart and Hallauer
(1968), Stuber and Moll (1970), Otsuka et al. (1972), and Stuber et al. (1973) have examined the relations among different types of hybrids. In most instances prediction methods based on single-cross performance were used and epistasis was proposed if predicted performance differed significantly from average performance of parental single crosses. In nearly all instances in all studies, significant epistatic effects were detected in specific instances. Two general conclusions are evident from all the studies:

1. Although bias from epistasis may be substantial in some unique combinations of lines, use of more complex procedures for predicting three-way and double crosses was not warranted.
2. Bias from genotype–environment interaction in prediction methods is greater than deviations from epistasis; hence adequate and extensive testing should be priority over inclusion of epistasis in prediction procedures. Dudley and Johnson (2009), however, proposed a prediction method that might be helpful at same number of testing environments.

Results of these studies indicate the presence of epistasis in crosses among elite inbred lines, but epistasis seems small in comparison with additive and dominance effects.

Russell and Eberhart (1970) and Russell (1970, 1976) used a method described by Fasoulas and Allard (1962) to estimate genetic effects of individual loci or short chromosome segments affecting quantitative traits (e.g., the classical polygene concept or the modern QTL concept and their effects). Their studies included backcross derived isogenic sub-lines of inbred lines B14 and Hy. Three gene loci that permitted identification of genotypes were used in their studies. Twenty-seven genotypes were included in each study and total variation among genotypes could be orthogonally partitioned into additive and dominance effects of each locus and epistatic effects among loci, which include all digenic and trigenic epistatic combinations. Analysis was on genotype means, permitting qualitative estimates of epistatic effects. Existence of epistatic effects was noted for nine traits in each study. In B14 for all traits except yield, additive, dominance, and epistatic effects accounted for 52.5, 6.2, and 41.2%, respectively, of total variability among genotypes. For yield per plant of B14, additive (31%) and dominance (28%) were nearly equal, with additive by additive (16%) and dominance by dominance (14%) epistasis also prominent. Additive effects (69%) were greater in the expression of yield of Hy; dominance effects (23%) were of secondary importance with all types of epistatic effects accounting for 8% of total yield variability among the 27 genotypes. These studies showed that epistatic effects were present for the three sets of loci considered in each inbred line. B14 and Hy are elite inbred lines that were used extensively after their release. Obviously, yield is affected by more than three loci and epistasis should be considered when designing long-term breeding procedures. Continuous improvement of economically important traits of maize hybrids will require identifying unique complexes of genes, which undoubtedly will include favorable epistatic combinations. Although estimates of components of epistatic
variances generally have not been successful for maize populations, mean comparisons of crosses produced among elite inbred lines invariably showed epistatic effects. In the first instance we are estimating variances from complex functions of covariances of relatives, often using either complex mating designs or combinations of mating designs. Large errors of estimates are inevitable. Mean comparisons are first-order statistics that are easier to estimate with smaller errors. In the first instance we are working with a population that includes a large collection of genotypes; and in the second we are comparing a few genotypes among elite inbred lines. As mentioned before, differences probably exist because of types of genetic materials studied and parameters estimated.

5.5 Correlations Among Traits and the Possibility for Indirect Selection

Genetic correlations are of interest to determine degree of association between traits and how they may enhance selection. Genetic correlations are useful if indirect selection gives greater response to selection for a trait than direct selection for the same trait. This depends on estimates of heritability for each trait and genetic correlation between them. Usually more than one trait is measured on progenies evaluated either for cyclical selection programs or in line development programs that require a combination of traits to satisfy growers. Although yield is usually the primary trait of interest, maturity, standability, grain quality, stalk quality, abiotic and biotic stress tolerance, etc. are all corollary traits that the modern maize breeder must consider for eventual usefulness of genotypes evaluated for grain yield. For instance, a high-yielding genotype will have limited use in the northern part of the US Corn Belt if it is very late flowering, has high grain moisture content at physiological maturity, and has a slow rate of moisture loss between physiological maturity and harvest maturity. Similarly, a high-yielding genotype that has very poor stalk quality is not acceptable in situations that use mechanical harvesting. It is, therefore, important to watch for desirable associations among traits during selection and testing of genotypes.

It was established in classical genetics that many genes have manifold effects; i.e., some genes seem to affect traits that are unrelated. Genes that have manifold effects are pleiotropic. In pleiotropy the same gene affects different traits in a complementary way; whereas in epistasis different genes affect the same trait. The existence of pleiotropic effects of genes in classical genetic analysis would logically imply the existence of pleiotropic effects for quantitative traits. Then it is possible that selection may be exerted on secondary traits that have greater heritability than the primary trait. For example, if genes with pleiotropic effects for kernel-row number and yield are present, selection on the basis of kernel-row number rather than yield itself may be more effective because kernel-row number has greater heritability (57.0% for kernel-row number vs. 18.7% for yield, Table 5.1) and it is easier to measure. Success of selection, however, also depends on association between both
traits. If the association is not large the effect of indirect selection for yield by use of kernel-row number will not be successful.

Indirect selection is the selection for a secondary character with the purpose to obtain a positive response in the desirable or primary trait. A high genetic correlation between the two traits and a high heritability on the secondary trait might provide genetic gain of the desired trait in less time and effort (e.g., prolificacy and grain yield, molecular markers and QTLs). In the past, morphological markers (visually screened in the field) such as flower color were used as secondary characters with the purpose to make progress on desirable traits. Indirect selection will be effective if

1. Heritability of the secondary trait is greater than that of the primary trait.
2. The genetic correlation between them is substantial.

Marker-assisted selection (MAS) has the same principles. Hence, a very high correlation is needed between the marker and the trait as well as very high heritability. Very few exceptions explain the effectiveness of indirect selection.

Mode and Robinson (1959) investigated the concept of genetic correlations for traits of maize under the assumption that genes exhibit pleiotropic effects. Linkage is another important cause of a correlation between traits, but a random mating population in linkage equilibrium was assumed. Genetic covariance of two traits could be partitioned in the same manner as genetic variance. It was possible, therefore, to conduct an analysis of covariance for two traits in the same fashion as an analysis of variance for each trait. Because expected mean cross-products had expectations similar to expected mean squares, it was possible to determine genetic and phenotypic correlation coefficients from analysis of mating designs proposed by Comstock and Robinson (1948). Many of the reports summarized in Table 5.1 also include estimates of genetic, additive genetic, and phenotypic correlations among traits. A few reports used the estimates to construct selection indices, but most merely reported associations among traits. Table 5.17 summarizes available estimates of genetic correlations among 13 traits of maize. All correlations were calculated from components of variances and covariances for different mating designs. Numerous estimates of genetic correlations between the yield and the plant and ear traits were available, and averages of estimates show, on an average, low association of yield with any trait.

In the example quoted previously for kernel-row number and yield, average genetic correlation is 0.24 (Table 5.17). Hence indirect selection (I) for yield improvement by use of kernel-row number would be less effective than direct selection for yield, assuming the same selection differentials for both traits:

\[ I = 0.24 \times \sqrt{57.0}/\sqrt{18.7} = 0.42 \]

Although the heritability of kernel-row number was 57.0% vs. 18.7% for yield, the genetic correlation (0.24) between kernel-row number and yield was small enough that indirect selection was 58% less effective than direct selection.

Kernel depth, ears per plant, ear length, and ear diameter had greater genetic correlations with yield than kernel-row number, but they also had smaller heritability.
### Table 5.17 Summary of genetic correlations among plant and ear traits with yield obtained by averaging the values across studies


a Number of estimates available for each pair of traits
estimates than kernel-row number. Indirect selection for yield improvement by these traits would be 36–46% less effective than selection for yield itself, which had an average heritability 50% smaller than the heritability estimates for ear traits (Table 5.1). Genetic correlations of traits with yield were, in all instances, too small to compensate for greater heritability estimates. Average genetic correlations with yield were larger for ear traits than for plant and ear height, days to flower, and tiller number. Kernel depth, for instance, had the highest correlation (0.51) with yield, but the coefficient of determination, that is the proportion of total sums of squares for yield that can be explained by linear regression on kernel depth, was only 26%. In all other instances, coefficients of determination with yield were 20% or less; and with a few exceptions, average correlations among plant and ear traits, including yield, were relatively small. Plant height and ear height had the greatest correlation \((r = 0.81)\), and some of the ear traits showed moderate associations. Some of the ear traits would be expected to have some association because of the type of measurements used. For example, cob diameter and kernel depth would be expected to show some association because both traits are included in diameter measurements of the ear. Kernel depth was positively correlated with kernel-row number (agreeing with the general observation that kernel depth increases as kernel-row number increases) but at the expense of kernel size (as evidenced by the negative correlation \((-0.33)\) between kernel-row number and kernel weight). Except for correlations of plant and ear height and kernel depth and ear diameter, coefficients of determination were less than 50%.

Estimates of genetic correlations for Iowa Stiff Stalk Synthetic (BSSS) are given in Table 5.18 for 10 traits. Average correlations given in Table 5.17 were obtained for several different populations. All correlations given in Table 5.18 do not have the same number of estimates included, and fewer estimates are included than for those in Table 5.17. Trends, however, are similar for the two sets of correlations.

Plant and ear height are highly correlated \((r = 0.77)\) and correlations among ear traits are relatively high. Correlation of kernels per row with yield is greater in Table 5.18, but only one estimate is available. The trait days to flower has a negative correlation with yield in Iowa Stiff Stalk Synthetic whereas the average of 13 estimates in Table 5.17 is a small, positive value. The two estimates for Iowa Stiff Stalk Synthetic were obtained from two sets of unselected inbred progenies that tend to show a negative relationship between flowering and yield. It is known that inbreeding produces a delay in flowering and, consequently, in seed set due to the availability of viable pollen.

Estimates of heritability for BSSS are given in Table 5.5 on a plot and progeny mean basis. The average of five estimates of the correlations between kernel depth and yield was 0.65, which was calculated from the analysis of variance and the covariance of progeny means. If we use average estimates of heritability for yield (41.4%) and kernel depth (54.8%) and assume the same selection differentials for each trait, we find that indirect selection for yield by use of kernel depth is only 74.8% as efficient as direct selection for yield. Only one estimate of the correlation of kernels per row with yield was available \((r = 0.84, \text{Table 5.18})\). If we use estimates of heritability reported by Bartual and Hallauer (1976), we find indirect
Table 5.18  Summary of genetic correlations among plant and ear traits with yield for BSSS

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>Ear</th>
<th>Cob Diameter</th>
<th>Kernel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yield</td>
<td>Plant</td>
<td>Ear</td>
<td>Length</td>
</tr>
<tr>
<td>Plant height</td>
<td>0.05 (5)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ear height</td>
<td>0.05 (5)</td>
<td>0.77 (4)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ear length</td>
<td>0.45 (5)</td>
<td>0.32 (3)</td>
<td>0.0 (3)</td>
<td>—</td>
</tr>
<tr>
<td>Ear diameter</td>
<td>0.54 (5)</td>
<td>—0.01 (3)</td>
<td>0.08 (3)</td>
<td>0.03 (5)</td>
</tr>
<tr>
<td>Cob diameter</td>
<td>0.13 (5)</td>
<td>0.19 (3)</td>
<td>0.23 (3)</td>
<td>0.15 (5)</td>
</tr>
<tr>
<td>Kernel-row no.</td>
<td>0.45 (2)</td>
<td>—</td>
<td>—</td>
<td>0.19 (2)</td>
</tr>
<tr>
<td>Kernels per row</td>
<td>0.84 (1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Kernel depth</td>
<td>0.65 (5)</td>
<td>0.32 (2)</td>
<td>—0.13 (3)</td>
<td>—0.06 (5)</td>
</tr>
<tr>
<td>Days to flower</td>
<td>−0.52 (2)</td>
<td>0.27 (1)</td>
<td>0.33 (1)</td>
<td>−0.14 (2)</td>
</tr>
</tbody>
</table>

a Number of estimates available for each pair of traits
Correlations Among Traits and the Possibility for Indirect Selection

5.5 Correlations Among Traits and the Possibility for Indirect Selection

selection for yield by use of kernels per row to be 15% less efficient than direct selection for yield. It would seem, in this instance, that indirect selection would be effective because of the high correlation between both traits and the high heritability of kernels per row (91%); however, the estimate of heritability for yield also was large (90%). If we use the average estimate of heritability (41.4%) for BSSS from Table 5.5, indirect selection would be more effective than direct selection for yield but it would be biased because of the different precision of the parameters used to calculate predicted response.

Genetic correlations inherently have large errors. All the estimates given in Tables 5.17 and 5.18 were obtained using components of variance and covariance calculated from the analyses of variance and the covariance. No simple correlations between, say, inbred lines and their F1 hybrids are even included; these correlations are found in Chapter 8. Precision of average estimates of genetic correlations among the traits in Table 5.17 is quite different because of the different number of estimates available. For example, 23 genetic correlations between plant and ear height were reported, whereas only three were available for tiller number and yield.

Few examples in the literature claimed indirect selection for yield improvement by one of the ear components to be successful. Among yield components, prolificacy (ears per plant) has been proposed as an important component for potential improvement of yield by indirect selection (Lonnquist, 1967; Mareck and Gardner, 1979; Singh et al., 1986; Subandi, 1990). Moreover, Zavala-Garcia et al. (1992) reported the importance of choosing environments to maximize genetic gain. This was based on the assumption that the best environment for improvement through selection would be the one where additive genetic variances were maximized. Plant density is an environmental factor that can be easily varied for more effective selection and genotypic and phenotypic correlations between yield and its components tended to be greater with increased plant densities (El-Lakany and Russell, 1971). However, selecting for prolificacy at low density (Coors and Mardones, 1989) and at two different density levels (Carena et al., 1998) did not increase grain yield. Also, selecting at the most favorable environment for trait expression (e.g., low plant densities) did not produce the best direct response to selection. Plant density affects selection response for grain yield and selecting at low plant densities might produce maize populations with improved ability to tolerate medium to high plant densities (Hyrkas and Carena, 2005). However, more evidence is desirable. Medium and high plant densities commonly used for selection might not be the best choice for breeders when screening early-generation lines across locations or for maximizing genetic improvement especially with heterogeneous genetic materials. Moreover, the ability of certain genotypes to tolerate high-density stress might not be associated with selection at high plant densities under certain selection methods during pre-breeding stages (Carena and Cross, 2003). If this is true, the breeder can increase the number of locations for testing due to less plot seed demands.

Another example was reported by Cortez-Mendoza (1977), who analyzed 10 generations of divergent mass selection for ear length. Direct selection for ear length was significant for both longer and shorter ears. Yield was not improved by selection for longer ears, but yield significantly decreased with selection for shorter ears. It was found that selection for increased ear length decreased kernel depth, which in
Table 5.18 had a greater correlation with yield than ear length. This trend was later confirmed after 27 and 30 cycles of selection for ear length (Lopez-Reynoso and Hallauer, 1997; Hallauer, 2005).

It seems that indirect selection for a complex trait, such as yield, is not plausible. Yield is an expression of fitness and drastic changes in one component of yield are accompanied by adjustments in other component(s), implying the existence of correlated changes of gene frequencies. It seems that the most effective method for yield improvement is direct selection for yield itself; there may be correlated changes among yield components, but these correlated changes will be in concert with development of the most physiologically efficient genotype for expression of yield. Ideally, yield should directly be included in an index (e.g. heritability index, rank-summation index, etc) that includes agronomic traits (e.g., lodging resistance, stress tolerance, grain quality, etc) that are economically important for the target region of cultivar development (e.g., breeding at the local target environments and not elsewhere).

The concept of linkage between simple visual genetic markers and quantitative traits was developed at the same time the inbred–hybrid concept was proposed in the public sector in the early 1900s. However, the efficacy of indirect selection with molecular markers depends, as in classical selection theory, on the genetic correlation between markers and traits and their respective heritability estimates (Johnson, 2004). This correlation is determined by the strength of linkage between the marker and the QTL.

References


Wright, S. 1922. The effects of inbreeding and crossbreeding on guinea pigs. III. Crosses between highly inbred families. USDA Bull. 1121:1–60.


Plant breeding has been defined by Nikolai I. Vavilov as plant evolution directed by man (Sanchez-Monge, 1993). Selection has been the essence of the overall science of plant breeding through the identification of elite germplasm and the combined application of methods available to the breeder. Evolution (via natural selection) and domestication (via artificial selection) created and improved the crop plant species that are so important for human survival. Ever since the potential of certain plant species as food sources was recognized, selection has been practiced for more productive plant types. Particularly in maize, in addition to great advances achieved by domestication and early empirical breeding, significant improvements have been made by changes in breeding methods that have occurred mainly during the past 100 years. New and old selection methods for the genetic improvement of maize still are important in increasing food production. Applied plant breeding programs and their targeted selection methods will allow the continuous production of more efficient cultivars which can maintain a sustainable production without the requirement of expensive inputs. Public breeding programs focus on not only short-term research goals but also long-term improvement of germplasm. Even though the product (e.g., hybrid, pure line) is the goal, some of the breeding strategies used for long-term selection are neglected. Long-term genetic improvement is needed for the success of short-term products. Future genetic gains are dependent on the deployment of useful genetic diversity carried out in the public sector (Smith, 2007).

Selection is ultimately the differential reproduction of genotypes and changes allele frequencies in the desired direction. The purpose and the critical feature of artificial selection are to choose from a group of individuals those that will be allowed to reproduce so as to make selection as effective as possible for a given selection intensity. Several techniques have been devised to help the breeder make accurate decisions. Such techniques involve procedures based on genetic concepts (such as progeny test or family evaluation) and experimental procedures that attempt to minimize the effect of environments on the expression of genotypes. In practical maize breeding, selection procedures can be divided into two distinct phases:
(1) Selection among populations (e.g., for breeding purposes and/or seed production).

(2) Selection of genotypes within a population (e.g., inbred line development and hybrid identification).

### 6.1 Selection Among Populations

In the first phase of the selection procedure breeders must decide which populations are more suitable for their purposes and resources. Neither breeding methods nor cutting edge technology will be useful if populations chosen are not well chosen. Selected populations can be used directly either for production or for breeding purposes. Eventually several selected populations can be combined to develop a unique base population for breeding purposes or two selected populations can be crossed and the first generation of hybrid population can be used directly for grain production.

The genetic potential of populations for breeding purposes may be evaluated simply by observation of their performance or by analysis of their pedigree, origin, and past selection records. Intrinsic genetic properties of populations can only be evaluated through genetic designs. In certain situations (e.g., complex traits difficult to measure) classical data generated by mating designs (see Chapter 4) could be complemented with information from molecular genetics (e.g., yield and testcross, diallel, and genotyping data, Barata and Carena, 2006; rate of dry down, doubled-haploids, genotyping, and North Carolina design II, Yang and Carena, 2008). Parameters of primary importance are the mean, the additive genetic variance and its relative magnitude, the coefficient of heritability, and additive genetic correlations among the most important traits. The mean of the trait that will be changed by selection is important because selection started at a lower level would require several cycles in a population with lower performance to attain a higher level. Also, the presence of useful genetic variance is a requirement to permit progress from selection. Other genetic information, whenever possible to obtain, may be useful as criteria for selection among populations. For instance, adding epistasis for prediction has also been useful (Dudley and Johnson, 2009).

The cross between populations also provides valuable information relative to heterosis and combining ability. When a set of varieties or populations are available, information can be obtained from use of the diallel cross analysis. The diallel cross analysis proposed by Gardner and Eberhart (1966) has been widely used for the evaluation of varieties because information about the performance of populations per se and in crosses is provided on the basis of variety effects $v_j$; total heterosis effects $h_{ij}$; and the components of heterosis $\bar{h}$ (average heterosis), $h_j$ (variety heterosis), and $s_{ij}$ (specific heterosis), as shown in Chapter 4. Diallel mating designs have lately been used in populations to identify alternative heterotic patterns (Melani and Carena, 2005), to identify competitive population hybrids and exploit heterotic combinations (Carena, 2005; Carena and Wicks III, 2006), to identify unique groups of genetic diversity for grain quality (Osorno and Carena, 2008) and its components,
and to identify maternal and reciprocal effects across populations for different traits (Jumbo and Carena, 2008).

The diallel cross also provides information that permits prediction of composite means allowing selection among populations (composites) that could be synthesized from a fixed set of varieties. Thus selection among populations is possible not only within a given set of varieties but also among all possible populations that could be synthesized from them by making use of prediction procedures (see Chapter 10 for details about prediction). Genetic variability within newly formed populations is difficult to predict but can be assured to a certain extent by combining genetically divergent varieties (see Fig. 5.1). The magnitude of heterosis is the most direct indication of amount of divergence among populations. Combining populations that give high heterosis when crossed would assure greater genetic variability within the newly formed population. However, Moll et al. (1965) pointed out that heterosis increases up to a certain level of divergence between populations, beyond which it tends to decrease. Lack of heterosis, however, is not always an indication of lack of genetic divergence between populations (Cress, 1966). Selection can influence the magnitude of heterosis, and the level of population improvement can have even more influence than how distant they are.

6.2 Selection of Genotypes Within Populations

In order to maximize the genetic improvement of populations, selection among genotypes is performed. In this case one can consider a population either as finite in size and consisting of a fixed set of defined genotypes or as a unique gene pool that for practical purposes is a random sample of a population theoretically infinite in size. In the first instance we may include all types of hybrids from inbred lines.

6.2.1 Selection Among Hybrids

Inbred lines drawn from a population may be considered a sample from a large population of genotypes. But once the lines (e.g., experimental lines) are drawn and crossed with lines from other sources (e.g., industry testers) for hybrid combination tests (e.g., in this case late or advanced generation hybrid trials), we have a fixed set of genotypes to consider. In this case selection would be among all possible hybrids involving a fixed set of parental lines (e.g., experimental lines crossed with several industry testers representing different heterotic groups). For example, with \( n \) inbred lines, selection could be done among \( n(n - 1)/2 \) single crosses and subsequently the best hybrids could be used directly for grain production or the top line could be released as parental or breeding sources. Hybrid development, however, is a somewhat static process because the objective is to obtain a genotype or a group of genotypes that, once identified, may be reproduced indefinitely without need for continuous selection. Actually the exact reproduction of hybrid genotypes is valid only for single crosses from completely homozygous lines, but
for practical purposes all other types of hybrids are assumed to be reproducible in
the same manner. Selection during the inbreeding process for line development has
a continuous phase that ends at the time when the lines are approximately homozy-
gous in contrast to the doubled haploid method of inbred line development in which
no selection is performed during inbreeding. Exceptional hybrids are not consid-
ered static entities when the parent lines are continuously improved through some
method like backcross or convergent improvement.

Cockerham (1961) gave evidence for the advantages of utilizing single-cross
hybrids, a current practice today. The author formalized the differences of selection
among single-, three-way, and double-cross hybrids. The expected mean squares
for the three types of relatives were expressed in terms of covariances of relatives
and translations made to genetic components of variance. Briefly, components of
 genetic variance show that selection among single-cross hybrids will always be
twice as effective as selection among double-cross hybrids, with selection among
three-way crosses intermediate to single and double crosses. The ratio of the addi-
tive genetic variance component was $1: \frac{3}{4}: \frac{1}{2}$ for single, three-way, and double
crosses, respectively. Also, Cockerham (1961) found that selection effectiveness
among single crosses will be even greater than among double crosses if non-additive
effects were important. A drawback from Cockerham’s study is that results are only
for hybrids produced from lines developed from one population. However, it seems
the results also should be applicable for groups of hybrids developed from lines
originating from different source populations. In most maize breeding programs,
inbred lines used to produce hybrids are from different source populations; the
object is usually to have the source populations as similar as possible within het-
erotic groups and as different as is feasible to enhance heterosis among crosses of
lines.

Theoretical approaches for the understanding of the inbred–hybrid breeding sys-
tem were also presented by Comstock (1964). Assume a random sample of inbred
lines from a non-inbred source population is crossed at random to produce single-
cross hybrids. Any hybrid produced in this manner will have a genotype that could
have occurred in the source population, and the probability of such a genotype in the
hybrid will be the same as the probability of that genotype in the source population.
Thus inbreeding followed by the crossing of inbred lines is not a method for creating
new genotypes and selection among hybrids is ultimately selection from the same
array of genotypes possible in the source population. In the same manner, when
hybrids are developed by crossing inbred lines from two different source populations,
the genotype of any such hybrid is one of the possible genotypes in the cross
between the two populations and occurs with the same probability. Advantages of
the inbred–hybrid system to select genotypes are listed below:

1. The selected genotypes can be reproduced and then submitted to extensive tests,
   allowing a greater precision of evaluation;
2. The commercial utilization of hybrids requires that they are reproducible, which
   is only possible through the inbred–hybrid system; and
3. Some processes used for inbred line development involve sequential selection.
In (3) inbred lines are eliminated at successive stages before the inbreeding is complete and thus, under effective selection, the probability of obtaining better genotypes is increased. This is not applicable for doubled haploid inbred line development process.

An expression to calculate theoretical gain by selection among hybrids obtained by crossing inbred lines from two source populations is given by Comstock (1964) as follows:

$$\Delta G = k(\hat{\sigma}_y^2 / \hat{\sigma}_{ph}) + v$$

where $k$ is a function of effective selection intensity; $\hat{\sigma}_y^2$ is the total genetic variance among hybrids, $\hat{\sigma}_{ph}$ is the phenotypic standard deviation of the performance means on which selection is based (e.g., precision of tests increases the gain from selection), and $v$ is the hybrid vigor in variety crosses. Although the theory is important in understanding the inbred–hybrid breeding system, it is difficult to adapt it in practical breeding for the following reasons:

1. Inbred-line development involves a sequential process of selection so that some lines are eliminated by their poor performance in the first generations of inbreeding, while others are eliminated later on the basis of general combining ability tests.
2. A vigorous line in one of the populations is desired to be used as seed parent and the most vigorous lines are not always those that give the best hybrids.
3. Success of an inbred line development program is not always measured by hybrids developed from that specific program because some lines have been shown to produce very good hybrids in combination with some preexisting lines or lines developed from other programs using other populations (e.g., some programs use industry testers as their counterpart).
4. Several attributes other than yield (e.g., multi-trait and multi-stage selection) are considered in the selection program so that the effect of selection on particular traits is difficult to predict.

Another important question about inbred–hybrid breeding pertains to the effectiveness of selection within new samples from the same population. Theoretically, there is no advantage in re-sampling the same population because hybrid genotypes obtained by crossing lines in the second sample will be expected to be the same as from the first sample. The key factor is that given successive samples of equal size drawn in the same way from the same population, the probability of containing the most extreme individual is the same for all samples. On the other hand, given any number of samples, there is a finite though small probability that if one more sample is taken it will be found to contain a more extreme individual than the most extreme in any preceding sample. This is only a principle in sampling theory but from the breeding standpoint one can conclude that after a reasonable number of hybrids have been screened the probable return per unit effort in screening more from the same
population is sharply decreased (Comstock, 1964) unless re-sampling is performed in early generations of inbreeding where sampling is often small.

The fact that hybrid genotypes occur with the same probability as the same genotypes in the source population brings about an obvious conclusion:

To increase the probability of obtaining better hybrids from a given population the most direct way is to improve the population itself. Also, new sampling from the same population is recommended only if the source population is being continuously improved in the interval from one sample to another (e.g., different selection cycles). The same is true when using molecular genetics in selection and breeding. Molecular markers are by-products of genomic research that can be used for indirect selection or in a combination of phenotypic trait expression and markers (Johnson, 2004). However, molecular markers would have to be identified for each cycle of selection which can delay the process of genetic improvement of genetically broad-based populations. On the other hand, as F2 populations are the most routinely used in public and commercial maize breeding programs, there is an opportunity for using markers on populations expected to be at maximum linkage disequilibrium. In both types of programs early-generation hybrid testing is performed to accurately predict the combining ability of inbred lines. Therefore, combining marker (e.g., several methods can provide marker score values) and phenotypic information seems useful to predict breeding values with and without the inclusion of epistasis (Johnson, 2004; Dudley and Johnson, 2009). Simulation studies predict improvements in selection response than need to be confirmed extensively. With actual data, Dudley and Johnson (2009) added epistasis in their prediction model used in lines and testcrosses. Their correlations of predicted and observed means were high enough to be used in breeding.

6.2.2 Recurrent Selection

Pre-breeding includes the introduction, adaptation, evaluation, and improvement of germplasm resources for use in breeding programs (Hallauer and Carena, 2009). Once unique germplasm is introduced and adapted to a particular environment (see Chapter 11) it can be improved by recurrent selection which develops improved germplasm resources that are either directly or indirectly used to develop new cultivars. Fig. 6.1 shows an example on how North Dakota (ND) released lines were developed (Carena and Wanner, 2009) which represents an integration of pre-breeding for germplasm improvement with cultivar development. Long-term established public breeding programs have developed unique inbred lines this way. It is important to note the presence of two types of yield trials: recurrent selection trials and early–and late–generation hybrid trials.

The future of extensive maize germplasm enhancement on a long-term basis is still uncertain. Even though the problem has been recognized the number of public maize breeding programs enhancing germplasm and developing cultivars to aid in the increase of genetic diversity of commercial cultivars continues to decline. Public maize breeding programs provide genetic diversity in reserve (Duvick, 1981),
breeding creativity, and an insurance against genetic vulnerability through continuous genetic improvement of elite genetically broad-based breeding populations (Carena and Wicks III, 2006). Future genetic gains are dependent on the deployment of useful genetic diversity carried out in the public sector (Smith, 2007).

The inbred–hybrid (East, 1908; Shull, 1909) and the population-hybrid (Carena, 2005; Carena and Wicks III, 2006) maize concepts were developed in the public sector. The most successful maize germplasm was Iowa Stiff Stalk Synthetic or BSSS (Sprague, 1946), a genetically broad-based population. B73 (Russell, 1972) was derived after five cycles of half-sib recurrent selection on BSSS (Fig. 6.2) and generated billions of dollars to the corn production industry. Even though the odds of developing successful public lines from genetically broad-based improved populations seem to be low, it only requires one to make significant impact (Hallauer and Carena, 2009).

The term “recurrent selection” was first suggested by Jenkins in 1940 (Hallauer et al., 1988) for long-term early-generation testing and recombination of S1 progenies. However, Hayes and Garber (1919) were credited as the first users of recurrent

Fig. 6.1  Flow diagram of breeding steps utilized to develop many North Dakota and Iowa lines

Fig. 6.2  Flow diagram of breeding steps utilized to develop B73 before intellectual property
selection. The suggestion of Jenkins (1940) was later modified by Hull in 1945 (Lonnquist, 1952) to describe a method to improve the specific combining ability of a heterozygous population with a tester. Recurrent selection was later redefined as a group of breeding procedures consisting of recurrent cycles of selection for outstanding genotypes with a specific purpose in a heterozygous population and the subsequent recombination of the selected portion of the population (Lonnquist, 1952).

The greatest virtue of recurrent selection methods is that they are designed

(1) to increase the frequency of favorable alleles for traits quantitatively inherited (e.g., low heritability) while
(2) maintaining genetic variability for continued genetic improvement (Jenkins, 1940; Hull, 1945; Comstock et al., 1949; Horner, 1956).

The general scheme for improving populations through recurrent selection is

\[
\text{POPULATION (C0)} \rightarrow \text{PROGENY PRODUCTION} \rightarrow \text{PROGENY EVALUATION} \rightarrow \text{ANOVA} \rightarrow \text{INTERMATE SELECTED PROGENIES} \rightarrow \text{IMPROVED POPULATION (C1)}
\]

In general, it takes 2–4 years per cycle of selection for improving grain yield. However, 1 year of progeny selection per cycle can be achieved for other traits (e.g., progeny evaluation and recombination during the same season for cold tolerance, Sezegen and Carena, 2009). The time per cycle depends on the breeding method utilized and the resources available (e.g., two winter nursery seasons for early-maturing genotypes, Carena et al., 2009a, b).

For instance, recurrent selection on Iowa Stiff Stalk Synthetic had several challenges before and after 1970.

\[
\text{Genetic gain (ΔG)} = 2–4\% \text{ cycle}^{-1}
\]

Less environments
More years per cycle
Hand harvest

1939

1970

Mechanization, more environments
Less lodging
Earlier

1990s (15 cycles)
Selection after 1970 included selection for less lodging and earliness in addition to selection for grain yield. Varieties that were only improved for grain yield did not perform as well for producing elite inbred lines.

Quantitative traits are dependent upon a large number of genes each having a relatively minor effect as compared with environmental effects (Fisher, 1918; Lønnquist, 1963). Quantitative traits have been explained by polygenes (Mather, 1941), quantitative trait loci (QTL) (Geldermann, 1975), and chromosome segments affecting the quantitative trait (Falconer and Mackay, 1996). Sometimes, the distinction between quantitative and qualitative traits has not been clear in the use of these concepts and the reader should be aware. While the methodology for detection of molecular markers associated with QTL for complex traits has been relatively easy, several limitations (e.g., validation, sample sizes, interactions) have challenged marker-assisted selection for cultivar development utilizing QTL-mapping procedures. After reviewing 20 years of QTL mapping experiments Bernardo (2008) concluded that exploiting QTL in selection has been difficult. In addition, Heffner et al. (2009) recognized that the use of marker-assisted selection for the improvement of quantitative traits has stagnated. However, significant industry investment in this technology is trying to improve the odds of moving chromosome segments found in parental sources of elite hybrids and increase the prediction of breeding values in F2 populations. Heffner et al. (2009) have suggested genome selection (Meuwissen et al., 2001) to circumvent the deficiencies of selection based upon QTL systems, incorporating all marker information to avoid biased marker effect estimates, although they recognized work remains to be done to validate it empirically. Bernardo and Yu (2007) have suggested the use of genome-wide selection as an alternative method based on the use of cheap and abundant molecular markers without being subjected to the identification of markers with significant effects. But again, these hypotheses remain to be tested. Ultimately, a combination of cost-effective classical and modern techniques focused on traits that are still challenging to measure will help increase our current breeding efficiency to detect superior genotypes in elite germplasm.

The goal of increasing the frequency of favorable alleles while maintaining genetic variability is achieved first by the evaluation of progenies in multiple environments and second by the recombination of a large enough sample of selected progenies (Hallauer, 1990). The creation of a family structure and progeny testing in multiple environments allows the separation of phenotypic variation caused by genetic effects from environmental effects. Furthermore, genetic variation among crosses can be partitioned in general and specific combining ability groups. Additive genetic effects and additive × additive epistatic genetic effects are associated with general combining ability while non-additive genetic effects including dominance and other types of epistatic genetic effects are associated with specific combining ability (Sprague and Tatum, 1942). As a consequence of the complexity encountered in multi-trait and multi-stage selection Hallauer (1999) suggested that molecular markers do not have a prominent role in recurrent selection programs even though these programs (not extensively available anymore) could be a source of actual data to the current simulation studies presented in genome-wide selection. Besides,
the target populations for recurrent selection programs are genetically broad-based, require large samples, and are selected usually within public breeding programs because of their potential to contribute useful germplasm (Hallauer, 1992). Ideally, obtaining a source population with high mean performance and high frequency of favorable alleles is desirable. The source population also should have low genetic variance at loci homozygous for favorable alleles and high genetic variance at loci lacking favorable alleles (Dudley, 1982).

Commercial maize breeders have used recurrent selection in a limited way (Weatherspoon, 1973; Good, 1990). There was an attempt to compare various recurrent selection methods with the pedigree method of developing inbred lines (Duvick, 1977; Good, 1990). However, the development of elite inbred lines depends on the improvement of germplasm sources that may include both genetically narrow and broad-based populations. An important similarity to note is that both methods use evaluation strategies (see Fig. 6.1). Both breeding methods need to be complementary and will only be successful if both are given the same importance (Hallauer, 1985). Recurrent selection studies in maize were developed to obtain improved sources of germplasm for the potential extraction of inbred lines and their possible use in hybrids. Continuous germplasm improvement is very valuable as a source of new and diverse maize hybrids. The industry still needs inbred lines that are new, different, and unrelated to existing lines (Weatherspoon, 1973) and the sustained future success of the hybrid industry depends on increasing the genetic potential of germplasm (Coors, 1999).

6.2.3 Recurrent Selection Within Populations

Selection within populations that are unique gene pools is directed toward the improvement of the gene pool itself, which is done by increasing the frequency of favorable alleles within the population. This is a dynamic process, since gene frequency is changed gradually following a recurrent selection procedure. Simultaneous selection in two populations via a reciprocal recurrent selection scheme is also included in this category. Recurrent selection procedures available to increase gene frequencies of favorable alleles may be divided into two main categories (Moll and Stuber, 1974):

(1) Intra-population improvement: recurrent selection methods are utilized within a population; the response per se is the direct response. An indirect response would be the one from the cross between a population and a tester.

(2) Inter-population improvement: recurrent selection methods are performed between two populations; the population cross is the direct response. An indirect response would be the response of populations per se.

Discussion about these two systems of selection is presented throughout this chapter. Improvement of a population as a unique entity may be done for qualitative traits
(usually controlled by one or few genes) or for quantitative traits (usually controlled by many genes), although the effect of selection is basically the same (i.e., change in allele frequencies). Additional information on selection theory can be found in Kempthorne (1957), Lerner (1958), Kearsey (1993), Moreno-Gonzalez and Cubero (1993), Arus and Moreno-Gonzalez (1993), Falconer and Mackay (1996), Lynch and Walsh (1998), and others.

Critical decisions to consider before embarking in recurrent selection programs are as follows:

1. **Choice of germplasm source**
   - Choice of germplasm is perhaps the most important decision to take. Need to review literature, visit stations, and choose elite populations (e.g., populations with high mean of the desired traits and adequate genetic variability).

2. **Produce progenies**
   - Choice of progenies (e.g., breeding method) depends on the crop, trait, and objectives.

3. **Evaluation of genotypes for selection**
   - The evaluation should try to minimize error (e.g., correct design, extensive testing) and should give as much information as possible. Analyses of variance provide variance component estimates of genotypic and environmental variance and a heritability estimate (or actually an estimate of repeatability) can be obtained on a progeny mean basis. Therefore, one can predict or actually measure how effective can be selection across recurrent selection methods.

4. **Recombination of selected genotypes**
   - We must equally intermate selected progenies by using diallel mating design (for small numbers), testcross isolations, bulk-entry method (for large numbers).

### 6.3 Intra-population Improvement: Qualitative Traits

Chapter 2 showed that gene frequency remains constant from generation to generation in a panmictic population in the absence of selection and other disturbing factors. Let us consider one locus with two alleles with frequency \( p(A) \) and \( q(a) \) in a maize population in Hardy–Weinberg equilibrium. The genotypic array is

\[
p^2(AA) : 2pq(Aa) : q^2(aa)
\]

Therefore, selection favoring either the dominant allele A or the recessive allele a will result in a change of allele and genotypic frequencies. Selection favoring recessives is the easiest one and can be performed fully in just one cycle of selection if penetrance or expressivity does not cause problems in the identification of recessive genotypes. For example, selection for sugary kernels is included in this category.
From a population segregating for sugary and normal kernels the breeder can easily obtain a completely sugary kernel population in only one cycle of selection.

Selection favoring recessives is common in maize breeding for several traits, such as sweetness, opacity, brachymy, lack of ligules. Most of them are completely recessive but even in these cases selection efficiency depends on how extensively the trait is influenced by environmental factors and genotypic background. For example, selection for brachytic plants in a heterogeneous population cannot usually be completed in only one generation if the breeder has interest in other traits and wants to keep effective population size as large as possible. The difficulty in selecting brachytic \((br_2br_2)\) plants is caused by the action of modifier genes that are polygenic in inheritance, resulting in a continuous distribution of phenotypes. So plants may vary from extremely small to nearly normal in height within the dwarf class. In the same way, other recessive genes may be affected by modifiers and then selection favoring recessives may be difficult.

If selection is against recessives another difficulty arises because recessive genes are retained in the heterozygous genotypes within the population. Although one can recognize homozygous recessives, they cannot be eliminated completely in only one generation of selection. If selection does not involve a progeny test, decrease in recessive homozygotes is gradual and asymptotic and rate of decrease depends on initial allele frequency. For the genotypic array \(p^2(AA):2pq(Aa):q^2(aa)\) and selection against the recessives, the reproductive rates are 1(AA):1(Aa):0(aa).

After selection the new genotypic proportion is

\[
[p^2/(p^2 + 2pq)]AA : [2pq/(p^2 + 2pq)]Aa
\]

The new gene frequencies are

\[
p_1(A) = 1/(p + 2q) \\
q_1(a) = q/(p + 2q) \\
p_1(A) + q_1(a) = 1.0
\]

Decrease in gene frequency is then

\[
\Delta q = q_0 - q_1 = q_0^2/(1 + q_0)
\]

where \(q_0\) is the initial gene frequency. For the complete elimination of recessive homozygotes the greatest selection response is when \(q_0\) is high (i.e., the greatest response when selecting favorable alleles is expected when \(q_0\) is high). If selection against recessives is continued, the new frequency in any generation can be related to the original allele frequency as

\[
q_n = q_0/(1 + nq_0)
\]

which gives an asymptotic pattern for change in allele frequency. Theoretically the limit of \(q_n\) is zero only after an infinite number of generations. However, one
can alternatively calculate the number $n$ of generations necessary to reduce the frequency of recessives from an initial frequency $q_0$ to a low frequency, say $q_n$, using

$$n = (q_0 - q_n)/(q_0q_n)$$

If selection does not completely eliminate the recessive homozygotes in each generation, the rate of change in allele frequency is lower than in the preceding case. If genotypes have a survival rate of $1(AA):1(Aa):(1 - s)(aa)$, where $s$ is the selection intensity for recessive homozygotes, the change in allele frequency is

$$\Delta q = -[sq^2(1 - q)]/(1 - sq^2)$$

Figure 6.3 shows the change in allele frequency $\Delta q$ for several levels of selection intensity and initial gene frequencies.

Selection can be quite effective for characters if the heterozygotes have an intermediate expression between the two homozygotes, as in the case of selection for yellow kernels. Most types of maize have colorless aleurone so that the kernel color is due to endosperm color. Endosperm is a triploid ($3n$) tissue and its genotype for color genes may be as follows:

<table>
<thead>
<tr>
<th>Endosperm genotype ($3n$)</th>
<th>Embryo genotype ($2n$)</th>
<th>Kernel color</th>
</tr>
</thead>
<tbody>
<tr>
<td>YYY</td>
<td>YY</td>
<td>orange</td>
</tr>
<tr>
<td>YYy</td>
<td>Yy</td>
<td>yellow</td>
</tr>
<tr>
<td>Yyy</td>
<td>Yy</td>
<td>light yellow</td>
</tr>
<tr>
<td>yyy</td>
<td>Yy</td>
<td>white</td>
</tr>
</tbody>
</table>

Careful selection for dark colored kernels increases the probability of eliminating heterozygous types and consequently leads to a rapid decrease in the frequency of the undesirable allele $y$. On the other hand, selection for white kernels may be more
efficient because they are easier to recognize among the four phenotypes for kernel color. The same could be applied to cob color when making inbreds more uniform (e.g., plant variety protection).

In practical maize breeding, selection at the one-gene level (as in the case of selection for yellow endosperm, aleurone color, brachysm, lack of ligules, cob color) may be done concurrently with recurrent selection programs where the greater emphasis is on quantitative traits. Some qualitative traits (such as opacity and sweetness) require special programs and some type of progeny test such as testcross or selfing is commonly used. Also selection in a backcross scheme is a very common procedure. Usually selection for qualitative traits is not completely dissociated from important quantitative characters such as yield, ear height, and lodging resistance.

### 6.4 Intra-population Improvement: Quantitative Traits

Selection for quantitative traits involves a quite different methodology from that used for qualitative traits because the breeder is concerned with a large number of traits with several genes and genotypes that cannot be individually classified. Quantitative traits involve some metrical measure and not the direct knowledge of genotypic or allele frequencies and the exact genetic constitution of individuals. Therefore, the tools used to study and measure selection effects must include statistical parameters such as means, variances, and covariances. Note, however, that change in allele and genotypic frequencies is the basic consequence of selection and changes in means and variances are the ultimate effect of those changes.

The breeder is first interested in the results of selection and its effect on a population mean. Theoretical and practical inferences about selection efficiency are drawn from manipulation of genetic and environmental parameters of genetic and environment reference populations.

#### 6.4.1 Response to Selection

Estimation of progress from selection has been one of the most important contributions of quantitative genetics to plant and animal breeding. One of its direct applications is concerned with the extent to which a given population is suitable for breeding purposes for either a given environment or a set of environments. Another important application is concerned with comparison of different selection methods. There are different approaches to determine theoretically the expected progress from selection. In any case a regression problem is involved.

For any breeding method, selection is based on a measurable entity (individual or family), which is the selection unit $X$ that is related to some individual $W$ in the improved population. Actually, $W$ is taken as an individual representative of all those related to $X$. In some cases, $W$ descends directly from $X$ but in most instances there is a recombination unit $R$ so that $W$ is related to $X$ through $R$ (i.e., the improved population does not descend directly from selected individuals).
6.4 Intra-population Improvement: Quantitative Traits

Expected progress from selection is related to the following basic question: If the best selection units are identified and recombined (directly or indirectly), what is the expected change in the population mean? When superior phenotypes are selected, it is assumed that superior genotypes are selected (i.e., phenotype and genotype are correlated to some extent). Otherwise, progress from selection would be impossible. The degree to which genotypic values of superior parents are transmitted to offspring depends on the heritability of the trait being selected. Considering only the linear relation between $X$ and $W$ for each unitary deviation in $X$, a response of $b_{WX}$ is expected in $W$, where $b_{WX}$ is the linear regression coefficient of $W$ on $X$ as follows:

$$b_{WX} = \frac{\hat{\text{Cov}}(W, X)}{\hat{\sigma}_X^2}$$

The denominator $\hat{\sigma}_X^2$ is the phenotypic variance of selection units. Denoting the mean of the selected group by $\bar{X}_s$, the deviation ($\bar{X}_s - \bar{X}$) is known as the selection differential $s$. So the first formula for the expected progress can be written as:

$$\Delta G = s \frac{\hat{\text{Cov}}(W, X)}{\hat{\sigma}_X^2}$$

The term $\hat{\text{Cov}}(W, X)$ is a type of covariance between relatives; in some cases it can be expressed as a linear function of components of genetic variance of the reference population (see Chapter 3). This feature simplifies the prediction procedure.

The selection differential is a measure of the difference between the mean of a subset $S$ containing the selected units and the mean of a set $B$ containing all units so that $S$ is contained in $B$, as represented in Fig. 6.4a. If truncation selection is used (Fig. 6.4b), the selection differential can be expressed in terms of phenotypic standard deviation units (i.e., $k = s/\hat{\sigma}_X$) and then it can be determined directly from properties of the normal distribution whenever the sample size (number of selection units) is greater than 50. Theory demonstrates that $k = z/p$, where $p$ is the proportion of selected units and $z$ is the height of the ordinate at the lower limit of the selected group. The value of $z$ can be obtained from Tables I and II of Fisher and Yates (1948) as follows:

![Fig. 6.4 Schematic presentation of (a) a subset $S$ selected from a basic set $B$ of entities and (b) its proportion $p$ in a normal distribution](image-url)
In Table I enter with $P$ where $P = 2p$ because the normal curve is symmetrical. Table I gives a relative deviate $x$ so that a proportion $1 - P$ is within the range from $-x$ to $+x$ and our value of $p$ is for only one-half the curve. The value of $x$ obtained in Table I is entered in Table II and $z$ is obtained directly. For example, if the top 10% is to be selected from a large sample, then $p = 0.10$ and $P = 0.20$. From Table I, we obtain $x = 1.28$, which we use to enter Table II, obtaining $z = 0.1758$. Thus the selection differential in standard units is $k = 0.1758/0.10 = 1.758$.

Alternatively, $k$ can be calculated with better approximation by obtaining $z$ directly from

$$z = [1/(2\pi)] e^{-X^2/2}$$

where $e = 2.718282$ and $\pi = 3.141593$. The value for $x$ is obtained from Table I of Fisher and Yates (1948). Values of $k$ for some of the more common selection intensities are given in Table 6.1.

**Table 6.1** Selection differential ($k$) in standard deviation units for several selection intensities ($p$) for sample size greater than 50

<table>
<thead>
<tr>
<th>$p$</th>
<th>$k$</th>
<th>$p$</th>
<th>$k$</th>
<th>$p$</th>
<th>$k$</th>
<th>$p$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>2.6652</td>
<td>0.11</td>
<td>1.7094</td>
<td>0.21</td>
<td>1.3724</td>
<td>0.31</td>
<td>1.1380</td>
</tr>
<tr>
<td>0.02</td>
<td>2.4209</td>
<td>0.12</td>
<td>1.6670</td>
<td>0.22</td>
<td>1.3459</td>
<td>0.32</td>
<td>1.1175</td>
</tr>
<tr>
<td>0.03</td>
<td>2.2681</td>
<td>0.13</td>
<td>1.6273</td>
<td>0.23</td>
<td>1.3202</td>
<td>0.33</td>
<td>1.0974</td>
</tr>
<tr>
<td>0.04</td>
<td>2.1543</td>
<td>0.14</td>
<td>1.5898</td>
<td>0.24</td>
<td>1.2953</td>
<td>0.34</td>
<td>1.0777</td>
</tr>
<tr>
<td>0.05</td>
<td>2.0627</td>
<td>0.15</td>
<td>1.5544</td>
<td>0.25</td>
<td>1.2711</td>
<td>0.35</td>
<td>1.0583</td>
</tr>
<tr>
<td>0.06</td>
<td>1.9854</td>
<td>0.16</td>
<td>1.5207</td>
<td>0.26</td>
<td>1.2476</td>
<td>0.36</td>
<td>1.0392</td>
</tr>
<tr>
<td>0.07</td>
<td>1.9181</td>
<td>0.17</td>
<td>1.4886</td>
<td>0.27</td>
<td>1.2246</td>
<td>0.37</td>
<td>1.0205</td>
</tr>
<tr>
<td>0.08</td>
<td>1.8583</td>
<td>0.18</td>
<td>1.4578</td>
<td>0.28</td>
<td>1.2022</td>
<td>0.38</td>
<td>1.0020</td>
</tr>
<tr>
<td>0.09</td>
<td>1.8043</td>
<td>0.19</td>
<td>1.4282</td>
<td>0.29</td>
<td>1.1804</td>
<td>0.39</td>
<td>0.9838</td>
</tr>
<tr>
<td>0.10</td>
<td>1.7550</td>
<td>0.20</td>
<td>1.3998</td>
<td>0.30</td>
<td>1.1590</td>
<td>0.40</td>
<td>0.9659</td>
</tr>
</tbody>
</table>

If the group from which selection is to be made has size $N \leq 50$, then $k$ must be obtained from Table XX (Fisher and Yates, 1948), where $N$ varies from 2 to 50. If $n$ (in the range from 1 to 25) is selected from a sample of $N$, the expected value of $k$ is the average of the first $n$ values in the column corresponding to $N$. For example, if 5 individuals are selected from a sample of 20, the expected $k$ value is the average of 1.87, 1.41, 1.13, 0.92, and 0.75 (i.e., $k = 1.216$). Table 6.2 gives some values of $k$ for the most common cases.

In the procedure so far only one selection unit is considered. Some breeding methods involve more than one selection unit, as for example in the case of among- and within-family selection. In such cases there is one expected gain for each selection unit; expected gains must be summed to give total gain. Also selection may be for only one or for both sexes. If selection is for both sexes sum the expected
gain for each sex to give total gain. At this point a general formula derived from multiple linear regression theory to predict the genetic gain from selection can be written as

$$
\Delta G = \sum_i \Delta G_i = \sum_i k_i c_i [\hat{\sigma}_A^2 (1 + \beta_i) / \hat{\sigma}_y^2]
$$

The subscript $i$ denotes either different selection units or different sexes; $\beta_i$ is a deviation from the additive genetic variance. When selection is not truncated, $k_i$ must be replaced by the selection differential $s$ in observed units and $\hat{\sigma}_y^2$ by $\hat{\sigma}_y^2$ (phenotypic variance). The coefficient $c_i$, which varies according to the breeding method, is the coefficient of transformation of covariance between relatives in components of genetic variance. When selection is for both sexes at the same selection intensity, summation results in $c$ being replaced by $2c$.

### 6.4.2 Intra-population Recurrent Selection Methods

The most common recurrent selection methods for maize intra-population improvement may be classified as follows:

1. Mass (phenotypic) selection
2. Family selection
   - (a) Half-sib family selection
   - (b) Full-sib family selection
   - (c) Selfed ($S_1$ or $S_2$) family selection (also known as inbred progeny selection)
   - (d) Combined selection
6.4.2.1 Mass (Phenotypic) Selection

For mass selection, individual plants are evaluated and selected phenotypically (i.e., no information other than their own phenotypes is used as a criterion for selection). Hence, the selection unit $X$ is the individual phenotype, which is also the recombination unit and the improved population $W$ is obtained directly from $X$ as follows:

\[ X \rightarrow W \]  
(selection unit) (improved population)

Usually the selected plants are pollinated by a random sample of pollen from the whole population (open pollination) so that selection is for only the female source of gametes, and there is no control for male gametes (pollen). The resemblance between $X$ and $W$ is due to a parent–offspring relationship and (see Chapter 3)

\[ \hat{\text{Cov}}(X, W) = \hat{\text{Cov}}(P - O) = \frac{1}{2} \hat{\sigma}_A^2 \]

Phenotypic variance among individual plants includes genetic variance among plants within the population and all types of environmental variance within the selection block. If epistasis is negligible, genetic variance is $\hat{\sigma}_A^2 + \hat{\sigma}_D^2$. Environmental variability can be expressed in terms of variance components if some experimental design is taken as reference, because mass selection is not based on replicated trials. For example, suppose the selection block is designed in a fashion similar to one replication of a randomized complete block design. Environmental variability within the block would be due to plot-to-plot environmental variance $\hat{\sigma}_p^2$ and within-plot environmental variance $\hat{\sigma}_{we}^2$. These components in the selection block are not expected to be exactly the same as those estimated over replicated trials. If we disregard this possible bias, phenotypic variance among individual plants is

\[ \hat{\sigma}_Y^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{we}^2 \]

Expected progress from selection for this form of mass selection can be written as

\[ \Delta G = \frac{k(\frac{1}{2})\hat{\sigma}_A^2}{(\hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{we}^2)^{\frac{1}{2}}} \]

Additive genetic variance can be estimated from mating designs given in Chapter 4. Phenotypic variance among individuals also can be estimated from a mating design where families are evaluated over replicated trials. For example, if half-sib families are taken as reference, phenotypic variance in the expected progress from selection for individual plants is (see Chapter 2)

\[ \hat{\sigma}_p^2 = \hat{\sigma}_f^2 + \hat{\sigma}_p^2 + (\hat{\sigma}_G^2 - \hat{\text{Cov}}\text{HS}) + \hat{\sigma}_{we}^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_p^2 + \hat{\sigma}_{we}^2 \]
Assume in the above example that the trait under mass selection is plant height and selected female parents are pollinated by the whole population (non-selected male gametes). Also assume that we conduct mass selection for the inverse situation, i.e., selection of male gametes but no control of female gametes. This is not a practical procedure but it could be done by covering female inflorescences (shoots) of all plants before flowering. At flowering time a sample of pollen is collected only from plants with the desired height and all plants are randomly pollinated with the bulk sample of pollen. The expected progress from selection would be, theoretically, the same as that given for female selection if the same selection intensity is used.

Selection for male gametes is feasible for traits that can be evaluated before flowering. For this situation, undesirable plants would be detasseled or eliminated so that selected plants would be pollinated by a sample of selected male gametes (provided by the same selected plants). An alternative procedure to select for both male and female gametes is by selfing the selected plants. In this case the selfed progeny must be recombined in an isolated block the following year. If recombination can be done in a winter nursery an additional year per cycle is not necessary. Selection for both sexes at the same selection intensity leads to the following expected progress:

\[
\Delta G = \frac{k\hat{\sigma}_A^2}{(\hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2)^{1/2}}
\]

Stratified mass selection was suggested by Gardner (1961), and the criterion for selection is the deviation of each individual phenotype from the mean phenotypic value of all plants within the same stratum. In this case, the isolated field is partitioned into sections and each section becomes the unit where selection is done. The differences among plants within sections are more likely to be due to genetic differences rather than due to environmental effects. Figure 6.5 represents an isolated field with 22,500 plants of a maize population that was adapted to North Dakota conditions (Eno and Carena, 2008). The field was subdivided into 16 sections (Fig. 6.5) in which the earliest 25 plants were selected. Even though selection was only done on the female, it was expected to have partial pollen control due to silks receiving pollen from earliest plants.

Hence,

\[
\text{Cov}(X, W) = \text{Cov}([Y_{ij} - \bar{Y}_i], W) = \text{Cov}(Y_{ij}, W) - \text{Cov}(\bar{Y}_i, W)
\]

where \( Y_{ij} \) is the \( j \)-th (\( j = 1, 2, \ldots, n \)) individual plant within the \( i \)-th (\( i = 1, 2, \ldots, s \)) stratum.

Taking the \( ij \)-th plant as the female plant from which \( W \) will descend

\[
\text{Cov}(Y_{ij}, W) = \text{Cov}(P - O) = (\hat{\nu}_2)\hat{\sigma}_A^2
\]

\[
\text{Cov}(\bar{Y}_i, W) = (1/n)\text{Cov}(P - O) = [1/(2n)]\hat{\sigma}_A^2
\]
Thus,

\[
\text{Cov}(X, W) = (\frac{1}{2})\hat{\sigma}_A^2 + \frac{1}{2n}\hat{\sigma}_A^2 \approx (\frac{1}{2})\hat{\sigma}_A^2
\]

if we consider \( n \) large enough that \( \frac{n-1}{n} \approx 1 \).

The phenotypic variance within strata over the whole population is

\[
\hat{\sigma}_y^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{wes}^2
\]

where \( \hat{\sigma}_{wes}^2 \) is the environmental variability within strata.

So the expected progress would be

\[
\Delta G = \frac{k(\frac{1}{2})\hat{\sigma}_A^2}{(\hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_{wes}^2)^{\frac{1}{2}}}
\]

which is doubled if selection is for both sexes at the same selection intensity. In the above formula, \( \hat{\sigma}_{wes}^2 \) should equal \( \hat{\sigma}_{we}^2 \) (previously defined) if strata and plots from replicated trials are of the same size and under the same environmental effects. If no stratification was used but the selection block was assumed to be hypothetically divided into strata of any size, the phenotypic variance includes the environmental variance among different strata:

\[
\hat{\sigma}_S^2, \text{ or } \hat{\sigma}_y^2 = \hat{\sigma}_A^2 + \hat{\sigma}_D^2 + \hat{\sigma}_S^2 + \hat{\sigma}_{we}^2
\]
Efficiency \((E)\) of stratification can be found by comparison of expected gain in each case:

\[
E = \left\{ \left( \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_S + \hat{\sigma}^2_{we} \right) / \left( \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_{we} \right) \right\}^{1/2} = \left\{ 1 + \left[ \hat{\sigma}^2_S / \left( \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_{we} \right) \right] \right\}^{1/2}
\]

The only purpose of this comparison is to show that the advantage of stratification increases for higher environmental variability among strata and stratification would show no advantage for \(\hat{\sigma}^2_S = 0\).

In all cases considered so far, selection is assumed for only one environment and all variance components are estimated for one environment. If selection is directed to a population of environments, appropriate estimates to predict genetic gain come from the population of environments. Gain is expected to be smaller than that for one environment if genotype–environment interaction is large. Mass selection is conducted in specific environments (e.g., locations), and genotype by environment interaction can only be estimated for a sample of environments. The phenotypic variance of the selection units is defined for one environment as

\[
\hat{\sigma}^2_y = \hat{\sigma}^2_A + \hat{\sigma}^2_D + \hat{\sigma}^2_p + \hat{\sigma}^2_{we}
\]

where \(\hat{\sigma}^2_A\) and \(\hat{\sigma}^2_D\) are biased estimates for one environment and \(\hat{\sigma}^2_A\) and \(\hat{\sigma}^2_D\) are for a population of environments. Expected gain for a random sample of environments is a direct function of \(\hat{\sigma}^2_A / \hat{\sigma}^2_y\).

Gardner (1961) estimated components of variance by use of the design I mating design in two locations; phenotypic variance was estimated by

\[
\hat{\sigma}^2_y = \hat{\sigma}^2_m + \hat{\sigma}^2_f + \hat{\sigma}^2_{me} + \hat{\sigma}^2_{fe} + \hat{\sigma}^2_p + \hat{\sigma}^2_w
\]

where \(\hat{\sigma}^2_m\) and \(\hat{\sigma}^2_f\) are the variances among males and females, respectively; \(\hat{\sigma}^2_{me}\) and \(\hat{\sigma}^2_{fe}\) are variances due to interaction of male and female effects by locations; \(\hat{\sigma}^2_p\) is the plot-to-plot environmental variance; and \(\hat{\sigma}^2_w\) is phenotypic variance within plots. The average realized gain in four generations was about 3.93% and the expected gain was 4.5%. Selection, as well as yield tests of selected materials, was conducted in one location. On the other hand, Hallauer and Sears (1969) obtained very little progress in two maize populations when selection was conducted in one location but yield tests for evaluation of progress were conducted at three different locations. Selection for simpler traits (e.g., maize flowering) in one environment, however, has been successful (Hallauer and Carena, 2009).

### 6.4.2.2 Family Selection

The primary difference between mass selection and family selection is that family selection is based on some type of progeny test (i.e., plant genotypes are evaluated on the basis of average performance of their progeny). The general scheme for family selection is as follows:
where 0 represents parental plants in the reference population, $X$ is the selection unit, $R$ is the recombination unit, and $W$ represents an individual in the improved population that is genetically related to $X$ through $R$ and $0$.

The important feature of family selection is that selection is based on family means (selection unit), which are obtained from replicated trials usually conducted over a set of environments. Family means, therefore, are expected to show a phenotypic variance smaller than that for individual plants. Also, genotype by environment interaction has a less pronounced effect on results from selection, as can be seen in the phenotypic variance expression shown below. However, the effect from genotype by environment interaction is affected by the amount of seed produced per progeny which is often limiting.

The model for a family value is

$$ Y_{ijkl} = m + f_i + b_j + E_k + (fE)_{ik} + e_{ij} + s_{ijkl} $$

where $m =$ mean, $f_i =$ families ($i = 1, 2, \ldots, p$), $b_j =$ replications/environments ($j = 1, 2, \ldots, r$), $E_k =$ environments ($k = 1, 2, \ldots, e$), $e_{ij} =$ experimental error associated with $i$th plot in $j$th replication, and $s_{ijkl} =$ plants/plots ($l = 1, 2, \ldots, n$).

Phenotypic variances among family means are

$$ \hat{\sigma}^2_{\gamma} = \hat{\sigma}^2_f + \hat{\sigma}^2_p/r + \hat{\sigma}^2_p/nr $$

for one environment, and

$$ \hat{\sigma}^2_{\gamma} = \hat{\sigma}^2_f + \hat{\sigma}^2_{fE}/e + \hat{\sigma}^2_p/(er) + \hat{\sigma}^2_w/(ern) $$

for more than one environment

where $\hat{\sigma}^2_f =$ variance (genetic) among family means for one environment, $\hat{\sigma}^2_{fE} =$ family by environment interaction variance, $\hat{\sigma}^2_p =$ plot-to-plot environmental variance (error variance), and $\hat{\sigma}^2_w =$ phenotypic variance within families. In either case, the expected progress from selection is

$$ \Delta G = kc(\hat{\sigma}^2_A + \beta_i)/\hat{\sigma}_{\gamma} $$
To facilitate the description of predicted progress we consider families grown in a experimental design in one environment. Genotype by environment interaction can be easily introduced when more than one environment is considered (see Chapter 4). Thus our discussion is based on the analysis of variance given in Table 6.3, in which estimates of all parameters are obtained on an individual plant basis. Often the analysis of variance is performed either on plot totals or on plot means and to express variance components on an individual plant basis the expectation of mean squares must be changed.

Table 6.3  Structure of the analysis of variance for families evaluated in a randomized complete block design in one environment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r−1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Families</td>
<td>f−1</td>
<td>M₁</td>
<td>(\hat{\sigma}^2_w + n\hat{\sigma}^2_p - +nr\hat{\sigma}^2_f)</td>
</tr>
<tr>
<td>Error</td>
<td>(r−1)(f−1)</td>
<td>M₂</td>
<td>(\hat{\sigma}^2_w + n\hat{\sigma}^2_p)</td>
</tr>
<tr>
<td>Within</td>
<td>rf(n−1)</td>
<td>M₃</td>
<td>(\hat{\sigma}^2_w)</td>
</tr>
</tbody>
</table>

(a) Half-Sib Family Selection

Half-Sib families can be used in intra-population improvement in several ways. We will introduce the following:

1. Modified ear-to-row
2. Modified–modified ear-to-row
3. Testcross

One commonly used half-sib procedure is that suggested by Lonnquist (1964) and known as “modified ear-to-row” selection. This method involves selection among and within half-sib families. Selection among families is based on half-sib family means that are compared with the population mean or the mean of all families. The selection unit is

\[ X = Y_{i,.} - \overline{Y}_{..} \]

where \(Y_{ijk}\) is the phenotype of the \(k\)th plant of the \(i\)th family in the \(j\)th replication.

A relationship of parent–offspring exists between \(R\) and \(W\) and \(R\) also is related as a half-sib with each of the \(nr\) plants of the \(i\)th family, so \(W\) is related as half-uncle-nephew with each plant in \(X\). Thus

\[ \text{Cov}(X, W) = \text{Cov}[(Y_{i,.} - \overline{Y}_{..}), W] = \left[1/(nr)\right] \sum_{jk} \text{Cov}(Y_{ijk}, W) = (1/8)\hat{\sigma}^2_A \]
Selection: Theory

because Cov(W, \bar{Y}) = 0. In this case, Cov(X, W) is half the genetic variance among half-sibs. The gain from selection among families is thus

\[ \Delta G_1 = k_1 (l_0) \hat{\sigma}_A^2 / \hat{\sigma}_p \]

In this case, \( \hat{\sigma}_f^2 = \hat{\sigma}_{HS}^2 \) is the genetic variance among half-sib families; it is also the covariance between half-sibs and can be used to estimate the additive genetic variance:

\[ \hat{\sigma}_f^2 = \text{Cov} \text{ HS} = (l_0) \hat{\sigma}_A^2 \]

However, selection experiments often are not large enough to provide good precision in genetic variance estimates.

The phenotypic variance \( \hat{\sigma}_p^2 \) among half-sib family means is

\[ \hat{\sigma}_{HS}^2 = \hat{\sigma}_f^2 + \hat{\sigma}_w^2 / (nr) \]

It can be estimated directly from the corresponding mean square by \( \hat{\sigma}_{HS}^2 = M_1 / (nr) \).

When half-sib families are evaluated over a series of environments the phenotypic variance is

\[ \hat{\sigma}_{HS}^2 = \hat{\sigma}_f^2 + \hat{\sigma}_{fe}^2 / e + \hat{\sigma}_w^2 / (er) + \hat{\sigma}_w^2 / (enr) \]

In the case of \( e = 1 \) (one environment), \( \hat{\sigma}_f^2 \) is overestimated because it actually estimates \( \hat{\sigma}_f^2 + \hat{\sigma}_{fe}^2 \), which causes an unknown bias in the expected gain. If the additive genetic variance is estimated as: \( \hat{\sigma}_A^2 = 4 \hat{\sigma}_f^2 \), then \( \hat{\sigma}_A^2 \) also is overestimated and the expected progress will be greater than the observed progress.

Within-family selection is conducted in the recombination block where all families are planted ear-to-row as females (which are detasseled) and pollinated by a sample of pollen produced by the inter-planted male rows (a bulk of all families). Selection is conducted only within the best families, but selected plants are pollinated by the whole population so that among-family selection is for only one sex. The within-family selection is essentially phenotypic and usually is also for only one sex. It can be represented by \( R \rightarrow W \), where \( R \) is one individual within a half-sib family (which is itself the recombination unit for among-family selection). The selection unit is the difference between each individual phenotype and the family mean to which it belongs. However, the family mean is observed in just one block, so the selection unit is \( \bar{Y}_{ijk} - \bar{Y}_{ij} \). A parent–offspring relationship exists between \( W \) and only one plant in the family; the remaining plants in the family are related to \( W \) as half-uncle-nephew. Hence:

\[ \text{Cov}(W, Y_{ijk} - \bar{Y}_{ij}) = \text{Cov}(W, Y_{ijk}) - \text{Cov}(W, \bar{Y}_{ij}) \]

But
\[
\begin{align*}
\hat{\text{Cov}}(W, Y_{ijk}) &= \hat{\text{Cov}}(P - O) = (\gamma_2)\hat{\sigma}_A^2 \\
\hat{\text{Cov}}(W, \bar{Y}_{ij}) &= (1/n)\hat{\text{Cov}}(P - O) + [(n - 1)/n]\hat{\text{Cov}}(\text{half - uncle - nephew}) \\
&= (1/n)((\gamma_2)\hat{\sigma}_A^2 + [(n - 1)/8]\hat{\sigma}_A^2) \\
&\approx (\gamma_8)\hat{\sigma}_A^2
\end{align*}
\]

when \( n \) is large. Thus

\[
\hat{\text{Cov}}(W, Y_{ijk} - \bar{Y}_{ij}) = (\gamma_2)\hat{\sigma}_A^2 - (\gamma_8)\hat{\sigma}_A^2 = (\gamma_8)\hat{\sigma}_A^2
\]

Gain from selection within families is given by

\[
\Delta G_2 = k_2 \frac{(\gamma_8)\hat{\sigma}_A^2}{\hat{\sigma}_w}
\]

where \( \hat{\sigma}_w^2 \) is the phenotypic variance within half-sib families. Total expected gain from selection among and within half-sib families is then

\[
\Delta G = \Delta G_1 + \Delta G_2 = k_1 \frac{(\gamma_8)\hat{\sigma}_A^2}{\hat{\sigma}_{HS}} + k_2 \frac{(\gamma_8)\hat{\sigma}_A^2}{\hat{\sigma}_w}
\]

under the assumption that genetic variance within the selected families is the same as that for the complete set of families.

An alternative procedure of modified ear-to-row selection is the recombination of only the best families after evaluation in replicated trials (Compton and Comstock, 1976), also known as “modified–modified ear-to-row” selection. In this case, remnant seed must be kept in good storage conditions for planting the next year. In the recombination block both female and male gametes come from selected families so that selection among families results in selection for both sexes. Within-family selection is also possible in this case, but it is usually done only on the basis of female phenotypes so that the expected gain is the same as in the preceding case. The total expected gain (Vencovsky, 1969) is thus

\[
\Delta G = \Delta G_1 + \Delta G_2 = c_1 \frac{(\gamma_8)\hat{\sigma}_A^2}{\hat{\sigma}_\text{HS}} + c_2 \frac{(\gamma_8)\hat{\sigma}_A^2}{\hat{\sigma}_w}
\]

An additional year per cycle is not necessary if recombination is done in off-season nurseries. Selection within families, however, is not recommended in this case. This scheme, involving only among-family selection, usually is referred to as “half-sib selection.”

The effect of genotype–environment interaction is the same as that discussed in the preceding case; i.e., if interaction is of great magnitude and families are evaluated over a series of environments, the additive genetic variance estimate tends to be smaller than that obtained for one environment. On the other hand, evaluation over environments has the advantages of decreasing the genotype–environment effect on phenotypic variance among family means and closing the gap between observed and predicted gain. For example, in a study reported by Paterniani et al. (1973), the effect of genotype by environment interaction was well illustrated.
Half-sib families were developed from an irradiated (with gamma rays) population and a control population and evaluated in 2 years. Additive genetic variance estimates were practically unchanged from one year to another for each subpopulation and were greater in magnitude for the irradiated one. However, combined analysis (both years) revealed a significant family by year interaction in the irradiated population leading to smaller estimates of $\hat{\sigma}_A^2$ and expected progress. No interaction was detected for the non-irradiated population. The results are summarized in Table 6.4.

Table 6.4 Additive genetic variance ($\hat{\sigma}_A^2$) and mean ($\bar{x}$) estimates and expected progress ($\Delta G$) in the population “Centralmex” (irradiated and control), using half-sib family selection (Paterniani et al., 1973)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Irradiated</td>
<td>1.2880</td>
<td>1.0624</td>
<td>0.1260</td>
<td>7.045</td>
<td>6.278</td>
<td>6.662</td>
<td>7.7</td>
<td>6.6</td>
<td>1.1</td>
</tr>
<tr>
<td>Control</td>
<td>0.7020</td>
<td>0.7184</td>
<td>0.7076</td>
<td>7.116</td>
<td>6.173</td>
<td>6.645</td>
<td>4.7</td>
<td>5.9</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Gain from selection among half-sib families can also be increased by making use of prolific plants so that one ear is open-pollinated and a second ear is selfed. Families are evaluated in replicated trials using seeds from open-pollinated ears (half-sib families) and recombination is performed using selfed seeds ($S_1$ families). If only the best $S_1$ families are recombined, selection is for both sexes and the expected gain is

$$\Delta G = k \frac{1/2 \hat{\sigma}_A^2}{\hat{\sigma}_{HS}}$$

where $\hat{\sigma}_{HS}$ is the square root of the phenotypic variance among half-sib family means (Empig et al., 1972). Under this selection scheme an additional year is necessary because recombination units are inbred families ($S_1$) and selfing results in a greater level of inbreeding. In the season following recombination the population is again non-inbred and can be used as source for new half-sib and $S_1$ families. Recombination can be performed in off-season nurseries because there is no selection within families, and this procedure reduces to 2 years per cycle instead of three. Expected progress does not take into account selection for prolificacy, but this may contribute to an additional gain from selection since prolificacy seems to be a moderately heritable trait (Lonnquist, 1967a; Hallauer, 1974; Carena et al., 1998).

The procedure reported by Lonnquist (1949) for the development of synthetic varieties also is a type of half-sib family selection. According to this procedure a series of plants is selfed and at the same time each male plant is crossed with a random sample of other plants of the population. The crosses are made in the same way as in design I described in Chapter 4. Each male plant, therefore, is evaluated in a number of families because individual plants are related as full-sibs within
families and as half-sibs among families. Selection among pollen parents is based on performance of the best progenies in replicated trials, and the new population is produced by intercrossing the inbred progenies of the selfed male parents. If the crossed families are grown according to design I structure, several components of variance can be estimated and the expected improvement is

$$\Delta G = \frac{k(\frac{1}{2})\hat{\sigma}_A^2}{\hat{\sigma}_m^2 + \hat{\sigma}_{me}^2 + [\hat{\sigma}_f^2 + \hat{\sigma}_{(f/m)e}^2]/n + \hat{\sigma}_p^2/r + \hat{\sigma}_w^2/(rk)}$$

where \(n\) is the number of plants to which each pollen parent was outcrossed, \(r\) is the number of replications, \(k\) is the number of plants per plot, \(\hat{\sigma}_m^2\) and \(\hat{\sigma}_{me}^2\) are variances due to males and females/males effects, and \(\hat{\sigma}_f^2\) and \(\hat{\sigma}_{r/m/e}^2\) are variances due to interaction of males and females/males effects by environments (Robinson et al., 1955; Table 4.23).

Half-sib family selection also can be used in a selection scheme known as the "testcross procedure." In this case the base population is crossed with a tester that can be either narrow base such as inbred lines or single crosses or broad base (including the parent population) such as varieties or populations. Families are evaluated in replicated trials, and selfed seeds from the best parents are used for recombination. Variances among tested half-sib families, as well as covariances between relatives, cannot be expressed as linear functions of genetic variances from the base population unless restrictions are imposed on either the gene frequencies or the genetic model. For this reason a different approach for the interpretation of genetic variances, covariances, and expected gain must be introduced. The genetic variance among crosses (\(\hat{\sigma}_f^2\) from Table 6.3) at the one-locus level (see Chapter 2) is

$$\hat{\sigma}_f^2 = (\frac{1}{2})pq[a + (s - r)d]^2$$

and the numerator for the expected progress is (Empig et al. 1972)

$$\hat{\text{Cov}}(X, W) = (\frac{1}{2})pq[a + (q - p)d][a + (s - r)d]$$

This is another instance, where \(\hat{\text{Cov}}(X, W)\) is not directly estimable from the estimated variance among testcrosses \(\hat{\sigma}_f^2\).

When the parent population is used as tester, \(p = r\) and \(q = s\), variance among families is \(\hat{\sigma}_f^2 = (\frac{1}{4})\hat{\sigma}_A^2\), and the numerator of expected progress is \(\hat{\text{Cov}}(X, W) = 2(\frac{1}{4})\hat{\sigma}_A^2\), which is exactly the same as that previously shown when selection is based on half-sib families and recombination from use of selfed seed. When only the best families are recombined the expected progress is

$$\Delta G = \frac{k(\frac{1}{2})\hat{\sigma}_A^2}{\hat{\sigma}_{\text{HS}}^2}$$
(b) Full-Sib Family Selection

Under this scheme selection units \( X \) are full-sib families evaluated in replicated trials and remnant full-sib seed is used to recombine the best families. Each cycle requires 2 years and new full-sib families are obtained from the improved selection cycle. If we consider selection for one sex,

\[
\hat{\text{Cov}}(X, W) = \hat{\text{Cov}}(\bar{Y}_{i.} - \bar{Y}_{..}, W) = \hat{\text{Cov}}(W, \bar{Y}_{i.}) - \hat{\text{Cov}}(W, \bar{Y}_{..})
\]

But

\[
\hat{\text{Cov}}(W, \bar{Y}_{i.}) = [1/(nr)] \sum_j \sum_k \hat{\text{Cov}}(W, Y_{ijk}) = (\frac{1}{4})\hat{\sigma}_A^2
\]

Because

\[
\hat{\text{Cov}}(W, Y_{ijk}) = \hat{\text{Cov}}(\text{uncle - nephew}) = (\frac{1}{4})\hat{\sigma}_A^2
\]

\[
\hat{\text{Cov}}(W, \bar{Y}_{..}) = 0
\]

The value of \( \hat{\text{Cov}}(X, W) \) is not directly estimable from the variance among full-sib families because

\[
\hat{\sigma}_f^2 = (\frac{1}{2})\hat{\sigma}_A^2 + (\frac{1}{4})\hat{\sigma}_D^2
\]

assuming no epistasis.

Expected progress from selection would be

\[
\Delta G = \frac{k(\frac{1}{4})\hat{\sigma}_A^2}{\hat{\sigma}_{FS}}
\]

where \( \hat{\sigma}_{FS} \) is the square root of the phenotypic variance among full-sib family means.

Usually selection is for both sexes \((c = 2 \times \frac{1}{4} = \frac{1}{2})\) because only the best families are intercrossed for recombination; for this case expected progress is

\[
\Delta G = \frac{k(\frac{1}{2})\hat{\sigma}_A^2}{\hat{\sigma}_{FS}}
\]

Use of prolific plants is also an alternative in full-sib family selection. Thus for each prolific plant one ear is used to produce full-sib families that are evaluated in replicated trials. A second ear for each parental plant is selfed and used for recombination of the best families. If selection is for both sexes, expected progress is the same as when remnant seeds from full-sib families are used for recombination. Hence there is no advantage in terms of expected progress if we recombine \( S_1 \) instead of full-sib families. Selection for prolificacy can result in an additional gain for yield that is, however, unpredictable.
(c) S₁ Family Selection

Under this scheme the selection units are S₁ family means compared with the grand mean of all S₁ families. Remnant seeds from the selfed ears are used for recombination. Since inbreeding is involved, \( \hat{C}ov(X, W) \) cannot be directly expressed as a linear function of variance components unless restrictions are imposed on either the gene frequency or the genetic model. For a one-locus situation we have

\[
\hat{C}ov(X, W) = 2pq[a + (q - p)d][a + (\frac{1}{2})(q - p)d] = \hat{\sigma}_{A_1}^2
\]

if selection is for both sexes.

The above expression can be represented as

\[
\hat{\sigma}_{A_1}^2 = \hat{\sigma}_A^2 + \beta_1
\]

where \( \beta_1 \) is a deviation from the additive genetic variance mainly due to dominance effects (Empig et al., 1972)

\[
\beta_1 = 2pq(p - \frac{1}{2})d[a + (q - p)d]
\]

Deviation \( \beta_1 = 0 \) for either \( p = q = 0.5 \) or a completely additive genetic model \( (d = 0 \text{ for all loci}) \). Thus the expected progress for S₁ family selection is

\[
\Delta G = k\hat{\sigma}_{A_1}^2 / \hat{\sigma}_{S_1}
\]

\[
= k(\hat{\sigma}_A^2 + \beta_1) / \hat{\sigma}_{S_1}
\]

where \( \hat{\sigma}_{S_1} \) is the square root of the phenotypic variance among S₁ family means.

The variance \( \hat{\sigma}_{A_1}^2 = \hat{\sigma}_A^2 + \beta_1 \) is not estimable from S₁ family trials because the variance among S₁ families has the following expectation (see Chapter 2):

\[
\hat{\sigma}_{A'}^2 + (\nu_4)\hat{\sigma}_{D}^2 = \hat{\sigma}_A^2 + \beta' + (\nu_4)\hat{\sigma}_{D}^2
\]

Hence expected progress can only be expressed in terms of additive genetic variance under the restriction of no dominance or gene frequency of 0.5.

(d) S₂ Family Selection

By this selection scheme, S₂ families are developed by selfing S₁ plants. Selection units are S₂ family means compared with the grand mean of all S₂ families. Recombination is with remnant seed from the selfed ears. As in the preceding case of S₁ family selection, \( \hat{C}ov(X, W) \) cannot be expressed as a linear function of variance components without a restriction on either the gene frequency or the genetic model. For one locus, we have
\[ \text{Cov}(X, W) = 3pq\sigma^2 + (\gamma_2)pq(q - p)ad + (\gamma_2)pq(q - p)^2d^2 \]

after multiplying by two to accommodate selection for both male and female gametes. Then, we can define the expression as

\[ \hat{\sigma}_{A2}^2 = (\gamma_2)\text{Cov}(X, W) \]

Then

\[ \text{Cov}(X, W) = (\gamma_2)\hat{\sigma}_{A2}^2 = (\gamma_2)(\hat{\sigma}_{A}^2 + \beta_2) \]

where \( \beta_2 \) is a deviation from additive genetic variance mainly due to dominance effects:

\[ \beta_2 = (10/3)pq(p - \frac{1}{2})d[a + (q - p)d] \]

Deviation \( \beta_2 = 0 \) either for \( p = q = 0.5 \) or for a completely genetic model. The expected progress is thus calculated as

\[ \Delta G = \frac{k(\gamma_2)\hat{\sigma}_{A2}^2}{\hat{\sigma}_{S2}} = \frac{k(\gamma_2)(\hat{\sigma}_{A}^2 + \beta_2)}{\hat{\sigma}_{S2}} \]

where \( \hat{\sigma}_{S2} \) is the square root of the phenotypic variance among \( S_2 \) family means.

As in the case of \( S_1 \) families, \( \hat{\sigma}_{A2}^2 = \hat{\sigma}_{A}^2 + \beta_2 \) is not estimable because the variance among \( S_2 \) family means is an estimate of \( (\gamma_2)(\hat{\sigma}_{A}^2 + \beta''') + (\gamma_16)\hat{\sigma}_{D}^2 \)

which equals \( (\gamma_2)\hat{\sigma}_{A}^2 \) only in a completely additive genetic model or with gene frequencies of 0.5.

In both \( S_1 \) and \( S_2 \) family selection, the number of years per selection cycle increases but with the advantage of having early-generation lines ready for inbred line development.

(e) Combined Selection

Combined selection makes use of two or more selection methods in the same program and can be classified into one of the following categories:

1. Those combining two or more methods alternatively.
2. Those using two or more methods simultaneously.

For example, the modified ear-to-row method of selection is included in the first category because the procedure alternates selection among families (progeny selection)
and within families (mass selection). Other combinations of selection methods are possible, e.g., combined S₁, mass, and reciprocal recurrent selection as proposed by Lonnquist (1967b, 1967c) or between S₁ and S₂ family selection to have additional pressure on S₁ progenies (see Chapter 12). Expected progress from a combination of selection methods is obtained by merely summing the expected progress for each individual method. In a broad sense most breeding methods would be considered as combined selection. For example, in all methods where prolific plants are required, phenotypic selection for prolificacy is done in addition to any further selection. In half-sib family selection, some mass selection is usually done when choosing parental plants. In addition to S₁ and S₂ generations, higher levels of inbreeding can be used in recurrent selection. When selected (elite) inbred lines Sₙ from one population are intermated to form a synthetic population, the process is similar to one cycle of recurrent selection with N generations per cycle (N > n). The effect of no planned extra selection is usually unpredictable but should be taken into account in evaluation of selection methods.

Simultaneous combinations of methods have not been studied comprehensively. Lonnquist and Castro (1967) suggested the use of combined selection through information obtained from simultaneous evaluation of half-sib and S₁ families, and results from this method have been reported (Goulas and Lonnquist, 1976). Briefly, the method includes obtaining half-sib and S₁ families from the same prolific plants. They are evaluated simultaneously, providing information for selection of superior individuals. If the criterion for selection (selection unit) is the mean of both types of families, and estimates of components of variance are available, expected progress can be calculated.

Family evaluation can be included in a split-plot design. If, for example, randomized complete blocks are used, each parental plant is evaluated in a whole plot (containing one half-sib and one S₁ family) and each family type is evaluated in a split plot. However, when half-sib and S₁ families are evaluated this way, competition between families can be a serious source of experimental error because of differences of vigor (due to inbreeding of S₁ families). Border rows can be used to reduce this source of error. An alternative procedure to avoid competition effects would be to keep the half-sib and the S₁ families in separate groups as a two-factor experiment with both factors in strips (Kempthorne, 1952). The analysis of variance is shown in Table 6.5.

From Table 6.5 we obtain \( \hat{\sigma}_p^2 = (M_6 - M_5)/(2r) \), which expresses the genetic variance among pairs of families. The phenotypic variance among pairs of families is

\[
\hat{\sigma}_F^2 = \hat{\sigma}_p^2 + \hat{\sigma}_a^2/r + \hat{\sigma}_c^2/(2r)
\]

which can be estimated by \( \hat{\sigma}_F^2 = M_6/(2r) \). Use of border rows is also recommended, but it would increase the size of the experiment.

Although families are evaluated in an experimental design similar to a split plot, separate analysis for each type of family can be obtained. If epistasis is negligible, an unbiased estimate of \( \hat{\sigma}_A^2 \) can be obtained from half-sib family analysis.
Table 6.5 Structure of the analysis of variance for half-sib (HS) and S1 family evaluation in an experiment with both factors in strips for one environment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS</th>
<th>E(MS)\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>$r-1$</td>
<td></td>
<td>$\hat{\sigma}<em>{c}^2 + 2\hat{\sigma}</em>{p}^2 + 2\hat{r}_{p}^2$</td>
</tr>
<tr>
<td>Parents (P)</td>
<td>$n-1$</td>
<td>$M_6$</td>
<td>$\hat{\sigma}<em>{c}^2 + 2\hat{\sigma}</em>{p}^2 + 2\hat{r}_{p}^2$</td>
</tr>
<tr>
<td>Error a (P × R)</td>
<td>($n-1$)($r-1$)</td>
<td>$M_5$</td>
<td>$\hat{\sigma}<em>{c}^2 + n \hat{\sigma}</em>{b}^2 + r \hat{\sigma}_{pf}^2 + nrK_f^2$</td>
</tr>
<tr>
<td>Families (F)</td>
<td>1</td>
<td>$M_4$</td>
<td>$\hat{\sigma}<em>{c}^2 + n \hat{\sigma}</em>{b}^2 + r \hat{\sigma}_{pf}^2$</td>
</tr>
<tr>
<td>Error b (F × R)</td>
<td>$r-1$</td>
<td>$M_3$</td>
<td>$\hat{\sigma}<em>{c}^2 + n \hat{\sigma}</em>{b}^2$</td>
</tr>
<tr>
<td>Interaction (P × F)</td>
<td>$n-1$</td>
<td>$M_2$</td>
<td>$\hat{\sigma}<em>{c}^2 + r \hat{\sigma}</em>{pf}^2$</td>
</tr>
<tr>
<td>Error c (P × F × R)</td>
<td>($n-1$)($r-1$)</td>
<td>$M_1$</td>
<td>$\hat{\sigma}_{c}^2$</td>
</tr>
</tbody>
</table>

\textsuperscript{a}$r$ is number of replications, $n$ is number of pairs of families, and $K_f^2 = \sum_k f_k^2$.

(\hat{\sigma}_A^2 = 4\hat{\sigma}_{HS}^2). If a completely additive model is assumed, \( \hat{\sigma}_A^2 \) can also be estimated from S1 family analysis (\( \hat{\sigma}_A^2 = \hat{\sigma}_{S1}^2 \)). If gene frequency is 0.5 for all segregating loci, \( \hat{\sigma}_S^2 \) can be estimated as \( \hat{\sigma}_A^2 + (\frac{1}{4})\hat{\sigma}_D^2 \) and an estimate of dominance variance can also be obtained.

The covariance between half-sib and S1 families can also estimate \( \hat{\sigma}_A^2 \) through the following relation:

\[ \widehat{\text{Cov}}(HS, S1) = (\frac{1}{4})\hat{\sigma}_A^2 \]

if epistasis is negligible and a completely additive model or gene frequency of 0.5 is assumed. If no evidence exists to support restrictions about either the genetic model or the gene frequency, an unbiased \( \hat{\sigma}_A^2 \) estimate (from half-sib families) is preferable.

The expected progress from selection depends on which type of family is used as the recombination unit to obtain the next generation. If the best half-sib families are recombined

\[ \widehat{\text{Cov}}(X, W) = (\frac{1}{4})\hat{\sigma}_A^2 + \beta_3 \]

where \( \beta_3 = (\frac{1}{4})pq(p - \frac{1}{2})d[a + (q - p)d] = 0 \) either for \( p = q = 0.5 \) or for a completely additive model. Thus

\[ \Delta G = \frac{k(\frac{1}{4})\hat{\sigma}_A^2 + \beta_3}{\hat{\sigma}_P} \]

If remnant seeds from S1 families are used for recombination

\[ \widehat{\text{Cov}}(X, W) = (\frac{1}{4})\hat{\sigma}_A^2 + \beta_3 \]
and the expected progress is given by

\[ \Delta G = \frac{k(3\hat{h}_A^2 + \beta_3)}{\hat{\sigma}_F^2} \]

where \( \hat{\sigma}_F \) is the square root of the phenotypic variance among family means. The numerator \( \hat{Cov}(X, W) \) corresponds to the average value of the two methods previously discussed: half-sib family selection using \( S_1 \) seeds for recombination and \( S_1 \) family selection considering either a gene frequency of 0.5 or an additive genetic model.

### 6.4.3 Inter-population Recurrent Selection Methods

Comstock et al. (1949) originally proposed reciprocal recurrent selection (RRS) to maximize use of both general and specific combining ability. Selection schemes designed to improve the cross between two populations are based on RRS (Table 6.6). All procedures have a common feature, i.e., improvement of populations by changing gene frequencies in a directed and complementary way so that a wide range of different types of gene action and interactions can be retained in the crossed population. Its most important feature is that selection is toward the improvement of populations themselves as well as the increase of heterosis in the crossed population. The performance of population hybrid progenies is the direct effect of selection. When both populations are improved in this way, their cross in the first generation can be used directly for grain production at a lower cost of seed production (Carena, 2005). However, maximization of heterotic effects can be attained only when improved populations are used as sources of inbred lines for hybrid development.

Improvement of parent populations as a source of inbred lines can also be effected through intra-population recurrent selection methods; but improvement

<table>
<thead>
<tr>
<th>RRS schemes(^a)</th>
<th>Selection unit(^b)</th>
<th>Recombination unit</th>
<th>Improved population ((A_1) and (B_1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Half-sib RRS</td>
<td>HS and FS</td>
<td>(S_1) families</td>
<td>progenies from (S_1 \times S_1)</td>
</tr>
<tr>
<td>2. Full-sib RRS</td>
<td>FS</td>
<td>(S_1) families</td>
<td>progenies from (S_1 \times S_1)</td>
</tr>
<tr>
<td>3. Modified</td>
<td>HS</td>
<td>(S_1) families</td>
<td>recombined HS</td>
</tr>
<tr>
<td>half-sib RRS-1</td>
<td></td>
<td>(S_1) families</td>
<td></td>
</tr>
<tr>
<td>4. Modified</td>
<td>HS</td>
<td>(S_1) families</td>
<td>recombined HS</td>
</tr>
<tr>
<td>half-sib RRS-2</td>
<td></td>
<td>(S_1) families</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Scheme 1: Comstock et al. (1949); scheme 2: Hallauer and Eberhart (1970); schemes 3 and 4: Paterniani (1967a, 1973, 1978) and Paterniani and Vencovsky (1977) respectively

\(^b\)HS: half-sibs; FS: full-sibs. In scheme 3, HS is a pooled half-sib family from a cross between an HS family from one population and the opposite population.
would be directed only to additive genetic effects and heterosis (due to non-additive effects) is changed only by chance. Crossing populations from different adaptation groups was suggested to improve hybrid response (Vasal et al., 1992). In addition, crosses between geographically isolated populations improved by intrapopulation selection programs can be productive as long-term genetic improvement is conducted (Carena and Wicks III, 2006).

In all instances crossed families are evaluated in replicated trials, and components of variance are obtained from the analysis of variance. If, for example, \( n_1 \) and \( n_2 \) families are obtained from populations A and B, respectively, and are evaluated in a randomized complete block design, estimates can be obtained from an analysis as shown in Table 6.7.

<table>
<thead>
<tr>
<th>SOV</th>
<th>Population A as male</th>
<th>Population B as male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>( M_a )</td>
<td>( M_B )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I + nf\hat{\sigma}^2_{b1} )</td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I + nf\hat{\sigma}^2_{b2} )</td>
</tr>
<tr>
<td>Families</td>
<td>( M_{1A} )</td>
<td>( M_{1B} )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I + nr\hat{\sigma}^2_{f12} )</td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I + nr\hat{\sigma}^2_{f21} )</td>
</tr>
<tr>
<td>Error</td>
<td>( M_{2A} )</td>
<td>( M_{2B} )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I )</td>
<td>( \hat{\sigma}^2_w + n\hat{\sigma}^2_I )</td>
</tr>
<tr>
<td>Within</td>
<td>( M_{3A} )</td>
<td>( M_{3B} )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma}^2_w )</td>
<td>( \hat{\sigma}^2_w )</td>
</tr>
</tbody>
</table>

Subscripts 1 and 2 denote populations A and B, respectively.

Following are the most used breeding methods for improvement of population crosses.

### 6.4.3.1 Half-Sib RRS

The selection procedure for half-sib RRS (Comstock et al., 1949) is based on families obtained from the cross between two populations, say \( A_0 \) and \( B_0 \). Crossed families are obtained in the same way as in design I described in Chapter 4 (i.e., plants in population \( A_0 \) are used as males and each is crossed with a number of plants in population \( B_0 \) used as females). Reciprocal crosses are made in the same way with different plants (e.g., isolations can be used). At the same time the plants used as males are selfed and their \( S_1 \) seeds are used as recombination units to obtain the improved populations (\( A_1 \) and \( B_1 \)). The procedure is restricted to open-pollinated varieties or populations with heterosis or representing heterotic patterns. Two sets of trials need to be managed so it is like managing two programs at the same time. The procedure is represented as follows:
6.4 Intra-population Improvement: Quantitative Traits

From Table 6.7, the following estimates can be determined:

\[ \hat{\sigma}_{f12}^2 = \frac{(M_{1A} - M_{2A})}{(nr)} \]
\[ \hat{\sigma}_{f12}^2 = \frac{(M_{1B} - M_{2B})}{(nr)} \]

which also estimate \( (\lambda_4)\hat{\sigma}_{A12}^2 \) and \( (\lambda_4)\hat{\sigma}_{A21}^2 \) (as shown in Chapter 2), respectively. Therefore, expected progress is obtained as follows:

\[ \Delta G = \frac{k_1\hat{\sigma}_{f12}^2}{\hat{\sigma}_{F12}} + \frac{k_2\hat{\sigma}_{f21}^2}{\hat{\sigma}_{F21}} = \frac{k_1(\lambda_4)\hat{\sigma}_{A12}^2}{\hat{\sigma}_{F12}} + \frac{k_2(\lambda_4)\hat{\sigma}_{A21}^2}{\hat{\sigma}_{F21}} \]

where \( \hat{\sigma}_{F12}^2 \) and \( \hat{\sigma}_{F21}^2 \), the square roots of the phenotypic variances among family means, can be estimated by \( M_{1A}/(nr) \) and \( M_{1B}/(nr) \), respectively. When more than one environment is considered the genotype–environment interaction is included.

If gene frequency is the same in both populations

\[ \hat{\sigma}_{A12}^2 = \hat{\sigma}_{A21}^2 = \hat{\sigma}_A^2 \]

Then the above expression will be the same as that given for intra-population recurrent selection improvement, where selection is based on half-sib families and \( S_1 \) seeds from the selected parents are used for recombination.
6.4.3.2 Full-Sib RRS (FR)

The techniques described for half-sib RRS are used for full-sib RRS except that full-sib progenies are evaluated instead of half-sib progenies, as described by Comstock et al. (1949). One main advantage is that twice as many plants can be sampled from the source populations to have the same number of progenies to evaluate as for half-sib RRS. Also, a major advantage is that only one set of trials is needed.

Full-sib RRS also has the main objective of improvement of the crossed population of the two base populations. Full-sib RRS is very effective if the two source populations produce two or more female inflorescences; i.e., one ear is used to produce the S\textsubscript{1} seed and the second ear is used to produce the full-sib seed (Hallauer and Eberhart, 1970). However, the production of S\textsubscript{1} progenies for making full-sib crosses is a good alternative for producing enough seed for extensive testing (Hallauer and Carena, 2009). The procedures of full-sib RRS can be illustrated as follows:

In Table 6.7, both $\hat{\sigma}_{f12}^2$ and $\hat{\sigma}_{f21}^2$ estimate the same variance (see Chapter 2):

$$ (\hat{\eta}_4)(\hat{\sigma}_{A12}^2 + \hat{\sigma}_{A21}^2) + (\hat{\eta}_4)\hat{\sigma}_{D12}^2 $$

Expected progress would be

$$ \Delta G = \frac{k(\hat{\eta}_4)(\hat{\sigma}_{A12}^2 + \hat{\sigma}_{A21}^2)}{\hat{\sigma}_F} $$

6.4.3.3 Modified RRS-1 (HS-RRS\textsubscript{1})

A modification of the original RRS was proposed by Paterniani (1967b). Briefly, the procedure is as follows. A number of open-pollinated ears (half-sib families)
are planted ear-to-row as females in a detasseling block where the male rows are plants from population B. In another isolated block, half-sib families from population B are used as females and male rows are from population A. From each block open-pollinated ears from half-sib families are harvested in bulk and evaluated in replicated trials. Remnant seed from half-sib families (not S1 progenies this time) from populations A and B is used for recombination to form A1 and B1, as follows (see Paterniani and Vencovsky, 1977):

The expected progress from selection is

\[ \Delta G = \frac{k_1(\bar{y}_{10})\hat{\sigma}_{A12}^2}{\hat{\sigma}_{F12}} + \frac{k_2(\bar{y}_{10})\hat{\sigma}_{A21}^2}{\hat{\sigma}_{F21}} \]

The modified method requires 3 years per cycle and it may be a cost-effective method to improve the population cross.

### 6.4.3.4 Modified RRS-2 (HS-RRS2)

Another modification of RRS proposed by Paterniani (1967b) makes use of prolific plants. Population A is planted in an isolated block I as females (detasseled) that are pollinated by population B (male rows). In another detasseling block II population B is used as female and population A as male. At flowering time a sample of pollen is collected from population A (male rows in block II) and used to pollinate the second ears of population A plants (female rows in block I). The reverse is done in block II, i.e., pollen collected from male rows in block I is used to pollinate female rows in block II. In each field the first ears in the female rows are open pollinated to form half-sib families (A plants × B population in block I and B plants × A population in block II). The crossed half-sib families are evaluated in replicated trials and recombination is done with hand-pollinated ears (half-sib families) from
each population. Therefore, the method is similar to RRS-1 but using prolific plants. The method is relatively simple and each cycle is completed in 2 years. Production of families and recombination are carried out in the same season. An expression that allows calculation of expected progress is given by Paterniani and Vencovsky (1978):

$$\Delta G = \frac{k_1(\hat{\sigma}_f^2)\hat{\sigma}_{A_{12}}^2}{\hat{\sigma}_{F_{12}}^2} + \frac{k_2(\hat{\sigma}_f^2)\hat{\sigma}_{A_{21}}^2}{\hat{\sigma}_{F_{21}}^2}$$

$$= \frac{k_1(\hat{\sigma}_f^2)\hat{\sigma}_{f_{12}}^2}{\hat{\sigma}_{F_{12}}^2} + \frac{k_2(\hat{\sigma}_f^2)\hat{\sigma}_{f_{21}}^2}{\hat{\sigma}_{F_{21}}^2}$$

where $\hat{\sigma}_{f_{12}}^2$ and $\hat{\sigma}_{f_{21}}^2$ are obtained from Table 6.7.

The analyses of variance in Tables 6.3 and 6.7 for intra- and inter-population family evaluation are based on single plant data and all components of variance are estimated in the same unit. Analyses of variance are commonly performed on either plot totals or plot means and the components of variance are estimated in their corresponding units. From analyses of variance based on either plot totals or plot means, components of variance can be estimated on an individual plant basis by changing the coefficients of the expected mean squares as shown in Table 6.8.

<table>
<thead>
<tr>
<th>E(MS)$^a$</th>
<th>SOV</th>
<th>Plot total</th>
<th>Plot mean</th>
<th>Individual plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Families</td>
<td>$n^2\hat{\sigma}_w^2 + n^2r\hat{\sigma}_l^2$</td>
<td>$\hat{\sigma}_w^2 + r\hat{\sigma}_l^2$</td>
<td>$n\hat{\sigma}_w^2 + nr\hat{\sigma}_l^2$</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>$n^2\hat{\sigma}_f^2$</td>
<td>$\hat{\sigma}_f^2$</td>
<td>$\hat{\sigma}_f^2$</td>
<td></td>
</tr>
</tbody>
</table>

$a$ is number of plants per plot; $r$ is number of replications; and $\hat{\sigma}_w^2 = \hat{\sigma}_f^2/n + \hat{\sigma}_f^2$, where $\hat{\sigma}_w^2$ is the within-plot variance and $\hat{\sigma}_f^2$ is the plot-to-plot environmental variance.

### 6.5 Comparing Breeding Methods

Maize breeding methods utilized to maximize germplasm improvement have evolved and have been modified by our increased knowledge in plant breeding and related sciences. Several procedures are available for the improvement of a population or a cross between two populations. Choice of the best procedure depends on the trait and population under selection, stage of the breeding program, and purpose of the breeding program. Chapter 12 gives more details on maize breeding plans and shows the peculiarities of each breeding and selection procedure.

At any stage of any breeding program, a breeder needs to know the effectiveness of a given selection procedure. Comparison of selection procedures relative to
their effectiveness by use of empirical results is a difficult task because several variables are involved. Breeders have used different methods on different populations under different environments and circumstances, usually with different objectives. Also, selection criteria are not expected to be the same among breeders. Sample size, selection intensity, and relative weights when selection is for several traits are examples of breeding procedures that differ among breeders. As shown in Chapter 7, however, several programs have been conducted specifically for comparing breeding methods.

Development of quantitative genetic theory has provided the basis for comparison of relative efficiency among several breeding systems. Because of difficulties already mentioned, theoretical comparisons are useful if estimated parameters are realistic. When comparing different selection methods, two approaches may be considered

(1) based on equal selection intensity
(2) based on equal effective population size.

The first approach is more important for comparing progress in short-term selection programs because the goal is to maximize gain within a few generations of selection. The second approach is considered if the selection program is intended to be long term; genetic variability in the population cannot be reduced drastically in a few generations if we are to expect continuous progress during the course of the program. When effective size of a population can be kept at a high level, however, differences in effective sizes resulting from different selection procedures are irrelevant. Rawlings (1970) pointed out that effective size usually is not a great problem in selection programs, effective population sizes of 30–45 should be a reasonable number. In spite of the potential presence of inbreeding depression and genetic drift, however, Guzman and Lamkey (2000) and Tabanao and Bernardo (2005) have shown no advantages in using large effective population sizes especially when non-additive genetic effects are present. Robertson (1960), on the other hand, emphasized the importance of effective size in selection programs because the total expected progress and the half-life of a recurrent selection program are proportional to effective population size. Robertson’s main conclusions were as follows:

(1) For a single gene with selective advantage \( s \), the chance of fixation (expected gene frequency at the limit after an infinite number of generations) is a function of the initial gene frequency and \( N_e s \), where \( N_e \) is the effective population size. Figure 6.3 gives the chance of fixation for genes acting additively. In artificial selection for a quantitative trait based on individual measurements, the selection differential can be expressed in \( k \) standard deviation \( \hat{\sigma} \) units and the coefficient of selection \( s \) can be found by \( s = 2ka/\hat{\sigma} \), where \( a \) is half the difference between the mean of the two homozygotes (on a metric scale). Therefore, for one gene with a given frequency and effect in a population the probability of fixation is a function of \( N_e k \) (Fig. 6.6).
The question of how long it would take to attain the selection limit is not adequate because the approach of gene frequency to the limit is asymptotic. However, it may be useful to know how long it would take for the mean gene frequency to get halfway to the limit or what would be the half-life of a selection process. The half-life for any selection process for additive genes will not be greater than $1.4N$ generations but it may be $2N_e$ for rare recessives. The half-life selection process, therefore, depends directly on the effective population size.

Few results have been reported on the probability of fixation for different selection procedures. Baker and Curnow (1969) examined the consequences of different effective population sizes on progress from selection within a population with a specific genetic model. The probability of fixation of a desirable allele for different initial gene frequencies is shown in Table 6.9, which shows that ultimate probability of fixation increases for greater effective population sizes.

Table 6.9 Probability of fixation (Pr) of the desirable allele in a population where initial gene frequency is $p$ and effective population size is $N_e$

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>Initial gene frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.1275</td>
</tr>
<tr>
<td>4</td>
<td>0.1584</td>
</tr>
<tr>
<td>16</td>
<td>0.3699</td>
</tr>
<tr>
<td>32</td>
<td>0.5982</td>
</tr>
<tr>
<td>64</td>
<td>0.8385</td>
</tr>
</tbody>
</table>

\[
Pr = \left( 1 - e^{-2Ne^p} \right) / \left( 1 - e^{-2Ne^s} \right) \] (Robertson, 1960)

Selection coefficient was given by Baker and Curnow (1969); $s = 0.28$ for selfing ($N_e = 1$) and $s = 0.14$ otherwise (selection of female plants with infinite number of males, i.e., random pollination).

Table 6.10 shows the increase in gene frequency over 1–10 generations for an allele with initial gene frequency $p = 0.2$, with dominance.

Baker and Curnow (1969) concluded that there would be a small difference in the frequency of the desirable homozygote when comparing effective sizes of 16 and 256 after some generations of selection, suggesting that a reasonably rapid progress from selection can be expected with small effective population sizes. In addition, a substantial added progress would be obtained if selection could be practiced within
6.5 Comparing Breeding Methods 263

Table 6.10 Gene frequency expected for 1, 5, 10, and infinite generations of selection under different effective population sizes for one dominant gene, with initial gene frequency $p = 0.2$ (Baker and Curnow, 1969)

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>0</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2</td>
<td>0.219</td>
<td>0.236</td>
<td>0.237</td>
<td>0.258</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.212</td>
<td>0.243</td>
<td>0.266</td>
<td>0.472</td>
</tr>
<tr>
<td>16</td>
<td>0.2</td>
<td>0.212</td>
<td>0.256</td>
<td>0.304</td>
<td>0.948</td>
</tr>
<tr>
<td>32</td>
<td>0.2</td>
<td>0.212</td>
<td>0.258</td>
<td>0.315</td>
<td>0.998</td>
</tr>
<tr>
<td>64</td>
<td>0.2</td>
<td>0.212</td>
<td>0.260</td>
<td>0.322</td>
<td>1.000</td>
</tr>
<tr>
<td>256</td>
<td>0.2</td>
<td>0.212</td>
<td>0.261</td>
<td>0.327</td>
<td>1.000</td>
</tr>
</tbody>
</table>

each of a number of replicate lines developed out of the same original population, followed by selection of the best replicate lines.

Vencovsky and Godoi (1976) used the expected change in gene frequency in one cycle of selection and the ultimate probability of fixation to investigate the relative power of three selection schemes combined with varying selection intensities:

(1) Selection among half-sib families: $\frac{5}{100}$ and $\frac{15}{100}$

(2) Selection among full-sib families: $\frac{10}{100}$ and $\frac{30}{100}$; and

(3) Mass selection: $\frac{25}{5,000}$ and $\frac{80}{16,000}$.

All the procedures lead to equivalence of effective population sizes. Using specific computer simulations for a large number of genetic and environmental parametric combinations, the following conclusions were obtained:

(a) Concerning the ultimate probability of fixation, mass selection was the most powerful method even for heritability as low as 0.05. Full-sib family selection generally ranked second. However, the selection coefficient $s$ was generally higher for the full-sib scheme, also indicating greater expected immediate response per cycle.

(b) When the expected gain per generation was measured, mass selection was the most efficient scheme and half-sib family selection ranked second for the selection intensities used. A lower selection pressure $\frac{500}{5,000}$ for mass selection was inefficient for immediate response, mainly under high genotype–environment interactions. Results suggest that selection programs should exploit mass selection primarily through higher selection pressure, which would provide a better balance between immediate response and long-range variability. This theory has been extensively proven with flowering time in maize (Hallauer and Carena, 2009).

Expected progress from selection has been the most widely used way to compare different selection methods, and the key formula is
\[
\Delta G = \left( \frac{k}{t} \right) [c_i(\hat{\sigma}^2_A + \beta_i)/\hat{\sigma}_Y]
\]

where \( k \) is a function of selection intensity, \( c_i \) is a coefficient that depends on the selection method and parental control, \( \beta_i \) is a deviation from additive genetic variance, \( \hat{\sigma}^2_A \) is the additive genetic variance, \( \hat{\sigma}_Y \) is the square root of the phenotypic variance of the selection unit, and \( t \) is the number of years required per cycle so that the expected gain is expressed as gain per year (Eberhart, 1970). Using the general formula, the first comparisons among methods for intra-population improvement are shown in Tables 6.11 and 6.12.

Table 6.11 Expected progress from several intra-population selection methods expressed by components of the general formula for one environment

<table>
<thead>
<tr>
<th>Method of selection</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
<th>( \hat{\sigma}^2_f )</th>
<th>( \hat{\sigma}^2_p )</th>
<th>( \hat{\sigma}^2_w )</th>
<th>( \beta )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass selection</td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Half-sib family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Modified ear-to-row</td>
<td>( \frac{1}{8} )</td>
<td>( \frac{1}{4} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within families</td>
<td></td>
<td></td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>2. Recombining S(_1)</td>
<td></td>
<td></td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>0</td>
</tr>
<tr>
<td>3. Testcross(c)</td>
<td></td>
<td></td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>—</td>
</tr>
<tr>
<td>Full-sib family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Recombining FS</td>
<td></td>
<td></td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>0</td>
</tr>
<tr>
<td>2. Recombining S(_1)</td>
<td></td>
<td></td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>0</td>
</tr>
<tr>
<td>Inbred family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. S(_1) family</td>
<td></td>
<td></td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>( \theta )</td>
<td>3</td>
</tr>
<tr>
<td>2. S(_2) family</td>
<td></td>
<td></td>
<td>( \frac{3}{2} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>( \frac{(\hat{\sigma}^2_A)}{2} )</td>
</tr>
<tr>
<td>Combined selection (HS+S(_1))</td>
<td></td>
<td></td>
<td>( \frac{3}{8} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>( \frac{(\hat{\sigma}^2_A)}{2} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( \frac{3}{8} )</td>
<td>1</td>
<td>( \frac{1}{r} )</td>
<td>( \frac{1}{kr} )</td>
<td>( \frac{(\hat{\sigma}^2_A)}{2} )</td>
</tr>
</tbody>
</table>

\( a \) \( c_1 \) and \( c_2 \) refer to selection for one and two sexes, respectively

\( b \) \( t = \) changes depending if recombination is conducted in the same year or the next year

\( c \) \( c_2 = \frac{1}{2} \) only for a completely additive model, when \( \beta = 0 \)

\( d \) \( \theta = 2pg(p - \frac{1}{2})d[a + (1 - 2p)d] \) at one locus level

\( e \) Substitute \( \hat{\sigma}^2_A \) for \( \hat{\sigma}^2 \) and \( \hat{\sigma}^2 \) for \( \hat{\sigma}^2_w \) (see Table 6.5)

Mass selection is the oldest and simplest of the schemes listed in Table 6.11. Its simplicity and the possibility of 1 cycle per year are the greatest advantages over other methods. The coefficient of \( \hat{\sigma}^2_A \) in the general formula is \( \frac{1}{2} \) (selection for one sex) or 1 (selection for both sexes), showing that the method makes use of a high proportion of additive genetic variance. Selection can be highly effective when the magnitude of genetic variance is large relative to non-genetic variances, as generally occurs with highly heritable traits. Its great disadvantage is that selection is based on individual phenotypes (1 on Table 6.12), resulting in a phenotypic variance among selection units greater than any other selection method. Field stratification, suggested by Gardner (1961), was intended to reduce the phenotypic variability and
Table 6.12 Coefficient (c) for additive genetic variance ($\hat{\sigma}_A^2$) relative to expected progress for several intra-population selection methods

<table>
<thead>
<tr>
<th>Selection method</th>
<th>Selection unit</th>
<th>Recombination unit</th>
<th>One sex</th>
<th>Both sexes$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mass</td>
<td>I</td>
<td>I</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>2. Mass</td>
<td>I</td>
<td>$S_1$</td>
<td>$\frac{3}{8}$</td>
<td>1</td>
</tr>
<tr>
<td>3. Modified ear-to-row HS</td>
<td>HS</td>
<td>$\frac{1}{8}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4. Modified ear-to-row HS</td>
<td>$S_1$</td>
<td>$\frac{1}{4}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5. Half-sib (testcross)$^c$ HT</td>
<td>$S_1$</td>
<td>$\frac{1}{4}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6. Full-sib</td>
<td>FS</td>
<td>FS</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>7. Full-sib</td>
<td>$S_1$</td>
<td>$\frac{1}{4}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8. Inbred ($S_1$)</td>
<td>$S_1$</td>
<td>$\frac{3}{4}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9. Inbred ($S_2$)</td>
<td>$S_2$</td>
<td>$\frac{3}{8}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10. Combined (HS-S1) HS</td>
<td>HS</td>
<td>$\frac{3}{16}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11. Combined (HS-S1) HS</td>
<td>$S_1$</td>
<td>$\frac{3}{8}$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Definition for additive genetic variance changes slightly with inbreeding and equals $\hat{\sigma}_A^2$ only for gene frequency one-half ($p = 0.5$) or no dominance ($d = 0$)

$^b$Assuming same selection intensity for both sexes

$^c$Testcross using a non-related population as tester

thus increase the expected progress. Therefore, it allows mass selection to be more effective for traits having lower heritability. Also, the possibility of selection for both sexes may increase the expected gain, as previously shown. One of the most important features of mass selection is that effective population size can be kept at a high level even if high selection pressures are used. This assures the breeder of maintaining genetic variability for several cycles and allows continuous gain from selection in long-term selection programs. The advantage of mass selection over family selection increases for characters with high heritability, such as flowering time in maize.

The advantage of any family selection procedure over mass selection is that genotypes are evaluated by means of a progeny test, i.e., families are evaluated in replicated trials. When experiments are replicated in a series of environments genotype–environment interaction is taken into account, which is not the case for mass selection. In all instances of family selection the phenotypic variance is reduced, thus increasing the expected gain. The coefficient of $\hat{\sigma}_A^2$, however, may be smaller than that for mass selection and the reduction in phenotypic variance must compensate for this difference. The advantage of family selection over mass selection increases for characters with low heritability, such as yield in maize. For half-sib family selection several procedures can be used. For selection among and within half-sib families and 1 year per cycle, the coefficient of $\hat{\sigma}_A^2$ is $\frac{1}{8}$ for among-family
and \( \frac{3}{8} \) for within-family selection (Table 6.12). The among-family coefficient can be increased to \( c = \frac{1}{4} \) if recombination is done in the next year using only remnant seed from the selected families, but the gain per cycle must be divided by two to be comparable with the preceding case \( (c = \frac{1}{8}) \) on a per year basis. Among-family selection assures a higher effective population size because all families rather than only selected ones are planted for recombination. When several populations are under selection in the same program, the second procedure permits the breeder to stagger yield trials and recombination phases among populations. Within-family selection is nearly as efficient as among-family selection in some instances (Webel and Lonnquist, 1967), but the relative gain in each phase depends on the ratio \( \hat{\sigma}_w^2 / \hat{\sigma}^2 \) and on selection intensity. As expected, relative efficiency of within-family selection increases for higher values of the heritability coefficient (Table 6.13).

Table 6.13 Expected gain in yield for selection among half-sib families in percent of total expected gain (among and within populations) for different combinations of selection intensities for three populations (Ramalho, 1977)

<table>
<thead>
<tr>
<th>Selection intensity (%)</th>
<th>Cateto</th>
<th>ESALQ-HV 1</th>
<th>Paulista Dent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among</td>
<td>Within</td>
<td>1-HS(^a)</td>
<td>2-HS(^a)</td>
</tr>
<tr>
<td>100</td>
<td>2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>49.8</td>
<td>66.5</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>53.6</td>
<td>69.8</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>64.3</td>
<td>78.3</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>75.3</td>
<td>85.9</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

\(^a\)1-HS and 2-HS refer to selection for one and for both sexes among families, respectively. In both cases, selection within families is for one sex.

Another alternative for selection among half-sib families is to recombine selfed seed obtained from a second ear in the parental plants. This procedure doubles the coefficient of \( \hat{\sigma}_A^2 (c = \frac{1}{2}) \), which not only increases the gain per cycle but also increases the number of seasons per cycle. If recombination is done in off-season nurseries, an additional year is not necessary. In addition, this procedure forces selection for prolificacy, which may result in an additional gain per cycle. Selection based on testcrosses by use of an unrelated population as tester is similar in structure to the preceding case. The only difference is that the variance among testcrosses (half-sib families) depends also on the genetic structure of the tester population. In the last two examples of half-sib family selection, effective population size is expected to be smaller than in any of the preceding cases because only selfed seed of the selected parents is represented in the recombination block. The expected gain increases when the average gene frequencies of the favorable alleles are low in the tester. In other words, it is desirable that gene frequencies \( r \) in the tester are low enough so that the quantity \( (p-r) \) would, on average, be positive and of significant magnitude (Comstock, 1964). However, in applied maize breeding
programs elite testers used in early-generation hybrid trials increase the prediction of late-generation hybrid trials when compared to poor testers (Hallauer and Carena, 2009).

The coefficient of $\hat{\sigma}^2_A (c = \frac{1}{2})$ is greater for full-sib family selection than for half-sib selection when remnant seeds of half-sib families are used for recombination. However, for fixed environments the phenotypic variance among full-sibs will also be greater than for half-sibs because the genetic portion of the phenotypic variance contains

$$
\left( \frac{nr + 1}{2nr} \right) \hat{\sigma}^2_A + \left( \frac{nr + 3}{4nr} \right) \hat{\sigma}^2_D \text{ for full sibs}
$$

$$
\left( \frac{nr + 3}{4nr} \right) \hat{\sigma}^2_A + \left( \frac{1}{nr} \right) \hat{\sigma}^2_D \text{ for half-sibs}
$$

with $n$ plants and $r$ replications. Effective population size is expected to be smaller for full-sibs than for half-sibs because only the parents involved in crosses are represented in the recombination block. The full-sib family procedure may also require more work because controlled pollination is involved, whereas mass selection and half-sib family selection can be accomplished by using only open-pollinated ears in isolated plantings except for the cases when $S_1$ seeds are used for recombination. However, the bulk-entry method allows for a number of rows that are just two times the effective population size.

Ramalho (1977) compared half-sib with full-sib family selection under realistic circumstances, using average estimates of genetic and environmental parameters. His comparisons were based on equality of selection intensity in the first instance and on equality of effective size in the second instance. Taking different combinations of selection intensities for among- and within-family selection, Ramalho (1977) showed that full-sib family selection tends to be more effective than half-sib selection (selection among and within half-sib families using remnant seed for recombination, i.e., two generations per cycle) for traits of lower heritability, which generally shows coefficients of variation of greater magnitude. All comparisons were based on a total selection intensity of 2%, combined for among- and within-family selection, as shown in Table 6.14. The relative effectiveness of half-sib and full-sib family selection for conditions specified in Table 6.14 is shown in Fig. 6.7.

An alternative for the recombination phase in the full-sib procedure is to use selfed seed from selected parents. The coefficient for $\hat{\sigma}^2_A$ in the expected gain, however, remains unchanged $(c = \frac{1}{2})$. Use of selfed seeds for recombination requires prolificacy in parental plants or an additional year for obtaining $S_1$ progenies. An additional year is also required unless recombination can be done in an off-season nursery.

Family selection using inbred progenies has the primary advantage of increasing genetic variability among families $(c = 1$ for $S_1$ and $c = \frac{3}{2}$ for $S_2$ families). Increase in number of years per cycle has caused these breeding systems to have limited use. If recombination can be done in an off-season nursery, number of years per
Table 6.14 Effective population size \( (N_e) \) for equal selection intensity for half-sib (HS) and full-sib (FS) family selection in relation to varying selection intensities for among and within half-sibs

<table>
<thead>
<tr>
<th>A. Selection intensity ( (i) ), %, for equal ( N_e )</th>
<th>B. Effective size for equality of ( i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-sib family</td>
<td>Full-sib family</td>
</tr>
<tr>
<td>Among</td>
<td>Within</td>
</tr>
<tr>
<td>(b) 10</td>
<td>20</td>
</tr>
<tr>
<td>(c) 20</td>
<td>10</td>
</tr>
<tr>
<td>(d) 25</td>
<td>8</td>
</tr>
</tbody>
</table>

Source: Adapted from Ramalho (1977)

**Fig. 6.7** Relative effectiveness of full-sib family selection (among and within) in percent half-sib family selection for varying values of heritability under the conditions: within-plot environmental variance, \( \hat{\sigma}^2_{we} / \hat{\sigma}^2_e = 7.0 \), and dominance variance/additive variance, \( \hat{\sigma}^2_D / \hat{\sigma}^2_A = 0.375 \). Upper (b): equal effective size; and lower (a): equal selection intensity as specified in Table 6.14. (CV is the experimental error coefficient of variation) (From Ramalho, 1977)

cycle can be reduced. In fact, if selection is for traits that are evaluated before flowering S1 progeny selection can be conducted at 1 cycle per year with the advantage of having elite early-generation lines already in the inbred-line development process. Effective population size is expected to be the smallest among the schemes so far discussed for a given number of selected progenies. Both S1 and S2 family selections are especially recommended for traits of very low heritability, so the increase in variance among families is of primary importance to allow some progress from selection. Use of inbred families also permits selection against undesirable recessive genes, thus providing a method for obtaining populations more suitable for the extraction of vigorous inbred lines.
Combined selection by use of information from half-sib and S1 progenies simultaneously makes use of a portion of additive genetic variance that is intermediate to those used by either type of progeny individually. When recombination is done using remnant seed from half-sib families the coefficient for $\hat{\sigma}_A^2$ is $\frac{3}{8}$, which is doubled ($\frac{3}{4}$) when S1 seeds are used for recombination. In the first case a higher effective population size is assured and only 2 years are required per cycle. In the second case an additional year or season is required. In all cases, however, the advantage of using information from both types of progenies is that different information relative to gene action is given by each type of progeny (Goulas and Lonnquist, 1976). Genes contributing to heterotic behavior are more likely to be selected in half-sib evaluation than in S1 family evaluation, where genes with favorable additive effects receive greater emphasis. Consequently, combined selection using information from both half-sib and S1 progenies provides for an increase in frequency of desirable alleles and allelic combinations more effectively than either half-sib or S1 separately.

A theoretical comparison among intra-population recurrent selection procedures was presented by Comstock (1964), based on their effectiveness in change of gene frequency in a population. The expressions for change in gene frequency were then transformed to relative magnitudes for a comparison of relative efficiency of selection systems. Such comparisons are shown in Table 6.15.

**Table 6.15 Expected change in gene frequency ($\Delta p$) per generation and relative values of selection systems for non-overdominant loci**

<table>
<thead>
<tr>
<th>Method</th>
<th>$\Delta p^a$</th>
<th>No dominance</th>
<th>Complete dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>$\frac{k}{2\hat{\sigma}_A}$</td>
<td>0.8–1.0</td>
<td>0.8–1.0</td>
</tr>
<tr>
<td>Full-sib family</td>
<td>$\frac{k}{4\hat{\sigma}_A}$</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Test progeny (1)$^b$</td>
<td>$\frac{k}{6\hat{\sigma}_A}$</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Test progeny (2)$^c$</td>
<td>$\frac{k}{6\hat{\sigma}_A} [A + 2(p - r)d]$</td>
<td>0.67</td>
<td>0.67–1.0</td>
</tr>
<tr>
<td>Selfed progeny</td>
<td>$\frac{k}{3\hat{\sigma}_A} [A + (p - \frac{1}{2})d]$</td>
<td>1.33</td>
<td>1.33–2.0</td>
</tr>
</tbody>
</table>

Source: Adapted from Comstock (1964)

$^aA = [a + (q - p)d]$, $^b$parental population used as tester; $^c$unrelated population used as tester

Formulas to predict selection gain from methods for inter-population improvement are shown in Table 6.16.

Few US temperate long-term reciprocal recurrent selection programs have been conducted in elite populations but some are under continuous progress since reciprocal recurrent selection programs have been especially productive. Coors (1999) showed that average annual gains from reciprocal recurrent selection programs exceeded
Table 6.16 Coefficient (c) for additive genetic variance (σ^2_A) relative to expected progress for inter-population selection methods

<table>
<thead>
<tr>
<th>Selection method^a</th>
<th>Selection unit</th>
<th>Recombination unit^b</th>
<th>One sex</th>
<th>Both sexes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(\hat{\sigma}^2_{A12})</td>
<td>(\hat{\sigma}^2_{A21})</td>
</tr>
<tr>
<td>1. HS-RRS</td>
<td>HS-FS</td>
<td>S_1</td>
<td>—</td>
<td>1/4</td>
</tr>
<tr>
<td>2. FS-RRS</td>
<td>FS</td>
<td>S_1</td>
<td>—</td>
<td>1/4</td>
</tr>
<tr>
<td>3. HS-RRS-1</td>
<td>HS</td>
<td>HS</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>4. HS-RRS-2</td>
<td>HS</td>
<td>HS</td>
<td>1/8</td>
<td>1/8</td>
</tr>
</tbody>
</table>

^aMethods 1, 3, and 4 have two components in the formula for expected progress. Method 2 has only one component and the numerator is \((\frac{1}{4})\hat{\sigma}^2_{A12} + (\frac{1}{4})\hat{\sigma}^2_{A21} = (\frac{1}{4})(\hat{\sigma}^2_{A12} + \hat{\sigma}^2_{A21}) = (\frac{1}{2})\hat{\sigma}^2_{A(12)}\).

^bSee Table 6.6.

significantly those achieved by breeders of commercial companies. Besides, interpopulation recurrent selection among early-maturing populations (Hallauer and Carena, 2009), tropical and subtropical populations (Beck et al., 1991; Vasal et al., 1992) and between tropical and temperate populations (Mickelson et al., 2001) are receiving greater emphasis.

The original RRS procedure proposed by Comstock et al. (1949) was effective (Collier, 1959; Penny et al., 1963; Moll and Stuber, 1971). Penny and Eberhart (1971) found very little improvement in the cross between two populations after four cycles of HS-RRS; but in a subsequent evaluation including five cycles (Eberhart et al., 1973), estimates of improvement were considerably greater than in the previous study. Despite the potential of the method, it has some inherent difficulties that may preclude its wide utilization. Some of these difficulties were outlined by Paterniani and Vencovsky (1977) as follows:

1. Selfing and crossing at the same time involve considerable labor, which reduces number of genotypes tested.
2. A sample of four or five female plants to be crossed with each male plant may not be adequate, thus reducing the accuracy of male evaluations. For this reason it should be more convenient to self the plants in a winter generation and make the crosses in the following year using a detasseling block, but this procedure increases the length of each cycle.
3. A more effective recombination among selected S_1 families may require an additional year (or winter) and thus the interval between cycles increases.
4. If each cycle interval becomes greater the gain is reduced when expressed on a per year basis. Tests are carried out in only 1 year, and selection is strongly affected by the season of the particular year. When cycle interval is shorter, the problems are mainly characterized by a less effective S_1 family recombination and by an inadequate number of females per male.

The methodology described by Penny and Eberhart (1971) includes 10 female plants for each male, which actually increases accuracy of genotype evaluation through
half-sib families. A less effective $S_1$ recombination is also a problem in FS-RRS, where each plant genotype is evaluated in a cross with another genotype. However, individual genotypes for each cross can be multiplied by selfing in each population to maintain the gametic array from each genotype. One advantage of this procedure is that high-performing families can be reproduced and inbred lines for hybrid development extracted from them through the procedure suggested by Hallauer (1967) and Lonnquist and Williams (1967). This combination makes FS-RRS an integrated method of population improvement and inbred line development. Options for a wider utilization of this method were described by Hallauer and Eberhart (1970).

A comparative study between HS-RRS and FS-RRS was reported by Jones et al. (1971), using theoretical comparisons of expected change in gene frequency and computer simulation. The relative efficiency of FS-RRS over HS-RRS is given by

$$\frac{\Delta G_{AF}}{\Delta G_{AH}} = \frac{k_F \hat{\sigma}_{PF}}{(k_H \hat{\sigma}_{PH})}$$

where $\Delta G_{AF}$ and $\Delta G_{AH}$ are the expected gene frequency changes in population A for FS-RRS and HS-RRS, respectively. Using realistic parameter estimates, it was found that

$$\frac{\hat{\sigma}_{PF}}{\hat{\sigma}_{PF}} = 1.18$$

Therefore, the selection differential for FS-RRS would need to be 1.2 times greater than for HS-RRS to give a similar response. Simulation results suggested that FS-RRS has an advantage over HS-RRS at lower selection intensities when the environmental variance is large relative to the total genetic variance. As selection intensity increases this advantage decreases.

Vencovsky (1977) made a comparison between FS-RRS and HS-RRS-2 relative to effective population size and expected progress. Effective population sizes for the FS-RRS and HS-RRS-2 are

$$N_e(\text{FS-RRS}) = \frac{2p_{FS} T}{(2 - p_{FS})}$$

$$N_e(\text{HS-RRS - 2}) = \frac{(8p_{HS} T)}{(4 + 3/M - 1/F)}$$

where $p_{FS}$ and $p_{HS}$ are the proportions of selected families in each method, $T$ is the total number of families tested, and $M$ and $F$ are the average number of males and females per male plant in half-sib families in the crossing block.

A comparison between the two formulas leads to the conclusion that to equalize the effective population size, $p_{FS}$ must be equal to $2p_{HS}/(p_{HS} + 0.25D)$, where $D$ is the denominator of the population size of HS-RRS-2. In terms of expected progress, relative efficiency would be

$$\text{HS-RRS - 2/FS-RRS} = 0.75k_{HS} \hat{\sigma}_{FS}/(k_{FS} \hat{\sigma}_{HS})$$

Using parameters approaching a realistic situation, values for $\hat{\sigma}_{FS}/\hat{\sigma}_{HS}$ were found to be in the range of 1.23–1.12, which is a good approximation to that by Jones
et al. (1971). Taking $\hat{\sigma}_{\text{FS}}/\hat{\sigma}_{\text{HS}} = 1.18$ and the proportion of selected families in HS-RRS-2 equal to 0.17, the relative efficiency is

$$\text{HS-RRS} - 2/\text{FS-RRS} = (0.75)(1.489)/[(1.159)(1.18)] = 1.137$$

for an approximation to equality of effective population size. This means that under the given conditions the expected progress for HS-RRS-2 would be 13.7% greater than that expected for FS-RRS. Such comparisons take into account differences in cycle duration, which would be, in this particular case, 3 years for FS-RRS and 2 years for HS-RSS-2. Another advantage of HS-RRS-2 over the two other methods is that larger samples are used in the crossing phase. When the comparison is based on equal selection intensities (but different effective population sizes) then

$$\Delta G_{\text{HS}}/\Delta G_{\text{FS}} = (0.75)(1.18) = 0.885$$

HS-RRS-2 would be 11.5% less effective than FS-RRS in changing the mean of the crossed population.

Paterniani and Vencovsky (1977) showed that reciprocal recurrent selection based on testcrosses of half-sib families (HS-RRS-1) makes use of a smaller portion of the additive genetic variance of the population cross than does the reciprocal recurrent selection based on prolific plants (HS-RRS-2). Consequently progress from selection is expected to be smaller for HS-RRS-1 but effective population size can be maintained at a higher level. It was shown for HS-RRS-1 that effective population size for each parental population is given by

$$N_e = 16S/(4 + 3/M - 1/F),$$

where $S$ is the number of half-sib families selected, $M$ is the number of remnant seeds taken per selected family for male rows, and $F$ is the number of detasseled plants per family from which ears are taken for the next cycle.

### 6.6 Increasing Gain from Selection

From the general prediction formula, there are several ways to increase the progress from selection and here we discuss them in some detail:

1. **Increasing selection pressure.** Selection differential ($k$) is a direct function of the proportion of selected units (i.e., selection intensity) so that $k$ increases as the selected group decreases in size, leading to a greater expected progress. However, selection intensity must be reasonably chosen because genetic variability can be reduced drastically with a very high selection pressure. If the program is long term, attention must be directed to the preservation of genetic variability, assuming a low but long-term gain from selection. A relative comparison of expected progress for short, intermediate, and long-term selection programs was shown in Fig. 1.3. Choosing appropriate selection intensities also depends on population size. Low selection intensity in small populations may lead to drastic changes in population structure due to small effective population size resulting in some inbreeding effect.
2. Adjusting the coefficient of $\hat{\sigma}_A^2$. The value of $c$ depends on the selection method used. For example, $c = \frac{1}{2}$ for mass selection and $c = \frac{1}{8}$ for half-sib family selection if only female gametes are selected. However, if both sexes are selected at the same selection intensity, then $c$ assumes values of 1 and $\frac{1}{4}$, respectively. Thus one can conclude that for any given method, parental control is an important way to increase progress from selection.

3. Increasing genetic variability. Genetic variability is determined at the time the population is either formed or chosen to undergo selection. Development of composite varieties from genetically divergent populations is very important as the first step in a breeding program (Eberhart et al., 1967). Once a population has been developed, the influence of $\hat{\sigma}_A^2$ on expected gain can be controlled only by selection (parental control) or by eventually limiting the range of environments.

4. Controlling environmental effects. Gain from selection increases directly by decreasing phenotypic variance among selection units. Considering a fixed amount of genetic variability, phenotypic variance can be decreased in several ways, most of them related to improvement of experimental techniques. For individual plant selection, phenotypic variance may be decreased by control of all factors that cause experimental error: (a) Choose uniform soils (in some soils, heterogeneity is an inherent property); topography; geographical conditions; or any factor that may limit normal growth, such as acidity, salinity, organic matter. In some special cases limited environmental conditions are representative of the environment for which the population is being selected. Also, cyclic crop rotation using uniform crops is a useful procedure to improve soil homogeneity. (b) Uniform soil preparation, fertilization, pesticide treatment (insecticides, herbicides, etc.), and other cultural practices should be used to minimize the differences because of cultural practices. (c) Care in data collection, in analysis, and in handling experimental material should be used. (d) Field stratification in small blocks and selection within blocks so that environmental variation among blocks is separable from phenotypic variance (among individuals within blocks) as suggested by Gardner (1961).

For family selection, an increase in number of replications and in number of plants per plot leads to a decrease in phenotypic variance and consequently an increase in genetic gains. Number of environments (locations or years, or both) also influences the magnitude of phenotypic variance. Such effects can be visualized in the following formula for phenotypic variance among family means:

$$\hat{\sigma}_F^2 = \hat{\sigma}_f^2 + \hat{\sigma}_{fe}^2/e + \hat{\sigma}_e^2/(re)$$

where $f$, $e$, and $r$ refer to number of families, environments, and replications, respectively. An increase in number of environments decreases $\hat{\sigma}_F^2$ because of reduction in $\hat{\sigma}_{fe}^2$ and $\hat{\sigma}_e^2$. Thus if selection is conducted for a population of environments, number of environments is the most powerful factor in reducing phenotypic variance. Expected gain will be affected by genotype–environment interactions within the defined population of environments. If only one environment is considered and selection is directed to a population of environments, expected response tends to
be overestimated. Usually the extensiveness of selection experiments is limited by space, amount of work, and available resources and facilities; an increase in number of environments usually must be at the expense of number of replications per environment. If genotype by environment interaction is of low magnitude (as known by previous studies or some other evidence), increase in $e$ at the expense of $r$ makes little difference in reduction of phenotypic variance. If number of replications is reduced to one, $\hat{\sigma}_e^2 + \hat{\sigma}_{IE}^2$ are confounded, leading to an unknown estimate of experimental error and reducing the precision of family mean estimates, particularly if genotype by environment interaction is large.

The effect of number of plants per plot in decreasing phenotypic variance has an asymptotic pattern and little change is obtained after a certain limit. For example, Eberhart (1970) found that very little increase in gain would be obtained with more than 15 or 20 plants per plot in two replications and four locations. But number of plants per plot cannot be decreased a great deal because sampling of genetic material is also of concern and may be an important source of experimental error. The increase in number of replications and locations also has a limit beyond which the gain in precision is very low (Eberhart, 1970).

Use of adequate experimental designs is another way to reduce phenotypic variability in family selection by removing some of the environmental variation. For example, lattice incomplete block designs may be used rather than randomized complete block designs when a large number of families are evaluated, as is usually done; lattice efficiency demonstrates that environmental variation exists within replications for several traits and can be removed to lower the experimental error.

Prediction of genetic gain depends on the accuracy of the heritability estimate, components of variance, and the relationship between real and expected responses (e.g., how large the genotype × environment interaction is). The selection differential (sometimes identified as $D$) depends on the proportion of individuals selected and the standard deviation ($\hat{\sigma}$) of the trait. The standard deviation of the trait represents the variability of the trait (e.g., how wide the frequency distribution is in Fig. 6.4) and it is expressed in terms of phenotypic standard deviation ($\hat{\sigma}_P$). The intensity of selection is the selection differential expressed in terms of phenotypic standard deviation (standardized selection differential or standardized selection coefficient, $k = D/\hat{\sigma}_P$).

Therefore, the selection differential is $D = k\hat{\sigma}_P$ and the expected response to selection is $\Delta G = k\hat{h}^2\hat{\sigma}_P$. Since $\hat{h}^2 = \hat{\sigma}_g^2/\hat{\sigma}_P^2$ then genetic gain becomes $\Delta G = k\hat{\sigma}_g^2/\hat{\sigma}_P$. Eberhart (1970) completed the formula as

$$\Delta G = \frac{k\hat{\sigma}_g^2}{y\hat{\sigma}_P}$$

being $y =$ years and $c =$ parental control. This allowed the comparison among different selection methods.

As seen before, the degree of parental control ($c$) depends on the family structure being used to intermate (recombination unit). The selection unit is the type of
family or individual plant being selected while the recombination unit is the type of family or individual being intercrossed with other selected genotypes to form the improved population. If we intermate half-sibs (half-sib is the recombination unit) \( c = 0.5 \) since selection is only being made among females. If selection is done for both males and females (e.g., resistant plants are selected before flowering, recombination unit is among full-sibs) then \( c = 1 \). When the selection unit is not the

<table>
<thead>
<tr>
<th>Methods of selection</th>
<th>Seasons per cycle</th>
<th>Parental control</th>
<th>( \hat{\sigma}_g^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( y )</td>
<td>( c )</td>
<td>( \hat{\sigma}_g^2 )</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma}_A^2 )</td>
<td>( \hat{\sigma}_D^2 )</td>
<td></td>
</tr>
<tr>
<td>Mass:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original</td>
<td>1</td>
<td>( \leq 0.5^b )</td>
<td>1.00</td>
</tr>
<tr>
<td>Selection after flowering (Gardner, 1961)</td>
<td>1</td>
<td>0.5</td>
<td>1.00</td>
</tr>
<tr>
<td>Selection before flowering</td>
<td>1</td>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Ear-to-row:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original (Hopkins, 1899)</td>
<td>1</td>
<td>( \leq 0.5^b )</td>
<td>0.25</td>
</tr>
<tr>
<td>Modified (Lonnquist, 1964)</td>
<td>1</td>
<td>0.5</td>
<td>0.25d</td>
</tr>
<tr>
<td>Modified–modified (Compton and Comstock, 1976)</td>
<td>2</td>
<td>1.0</td>
<td>0.25d</td>
</tr>
<tr>
<td>Half-sib families:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population as tester (Jenkins, 1940)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remnant half-sib seed</td>
<td>2</td>
<td>1.0</td>
<td>0.25</td>
</tr>
<tr>
<td>Self-pollinated seed</td>
<td>2</td>
<td>2.0</td>
<td>0.25+</td>
</tr>
<tr>
<td>Poor inbred of population</td>
<td>2</td>
<td>2.0</td>
<td>0.25+</td>
</tr>
<tr>
<td>Full-sib families:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant-to-plant crosses</td>
<td>2</td>
<td>1.0</td>
<td>0.50</td>
</tr>
<tr>
<td>Self-pollinated seed</td>
<td>3</td>
<td>2.0</td>
<td>0.50</td>
</tr>
<tr>
<td>Selfed progeny:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 (Hull, 1945)</td>
<td>2</td>
<td>1.0</td>
<td>1.00</td>
</tr>
<tr>
<td>S2</td>
<td>3</td>
<td>1.0</td>
<td>1.50</td>
</tr>
<tr>
<td>( S_n )</td>
<td>( n + 1 )</td>
<td>1.0</td>
<td>( \sim 2.00^e )</td>
</tr>
<tr>
<td>( S_1 ) modifieda (Dhillon and Khera, 1989)</td>
<td>1</td>
<td>0.5</td>
<td>( \sim 0.00^e )</td>
</tr>
<tr>
<td>Reciprocal recurrent selection:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Half-sib (Comstock et al., 1949)</td>
<td>2</td>
<td>1.0</td>
<td>0.25</td>
</tr>
<tr>
<td>Modified-1 (Paterniani and Vencovsky, 1977)</td>
<td>3</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Modified-2 (Paterniani and Vencovsky, 1977)</td>
<td>2</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>Modified-3 (Russell and Eberhart, 1975)</td>
<td>2</td>
<td>2.0</td>
<td>0.25</td>
</tr>
<tr>
<td>Full-sib (Hallauer and Eberhart, 1970)</td>
<td>2</td>
<td>1.0</td>
<td>0.50</td>
</tr>
<tr>
<td>Modified (Marquez-Sanchez, 1982)</td>
<td>3</td>
<td>0.5</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Source: Adapted from Hallauer and Carena (2009)

*The modified \( S_1 \) scheme also includes testcross phase
*Parental control equals 0.5 if adequate isolation was used
*Predicted gains were based on the formula suggested by Eberhart (1970) as \( \Delta G = (ck\hat{\sigma}_g^2)/(\hat{\sigma}_P) \)
*If within-plot plant selection was used then 0.375 \( \hat{\sigma}_A^2 \) would be added to predicted gain
*Depending on the level of inbreeding at which progenies are evaluated but coefficients approach 2 and 0 as \( F \) approaches 1 for single-seed descent with no selection
same as the recombination unit then \( c = 2 \) (e.g., half-sib selection when selfed seed of selected genotypes are used for recombination.

The genetic gain formula can help predict the breeding method that will be most effective for a particular situation. However, selection response can then be improved by increasing heritability and the intensity of selection. We have seen that environmental variation can be reduced using appropriate experimental designs, increasing the number of locations and, therefore, increasing heritability. We have also mentioned that we can increase population size in order to increase the intensity of selection for a trait. However, time, labor, and experimental facilities set the upper limit for population size. Also, inbreeding or fixation of alleles due to small effective population sizes set the lower limit. Genetic variation is maintained by recombining a large enough number of individuals to prevent and/or reduce genetic drift (e.g., random changes in allelic frequency different from expected). Table 6.17 shows the relative values for commonly used intra-population recurrent selection methods including the variance due to dominance deviations.

### 6.7 Correlation Between Traits and Correlated Response to Selection

Correlation, measured by a correlation coefficient, is important in plant breeding because it measures the degree of association, genetic or non-genetic, between two or more characters. If genetic association exists, selection for one trait will cause changes in other traits, basically a *correlated response*. The cause of correlation can be genetic and/or environmental. Genetic causes may be attributed to pleiotropism and/or linkage disequilibrium. Pleiotropism occurs when one gene affects simultaneously several physiological pathways, resulting in influence over several observed characters. Linkage disequilibrium refers to genes which show a tendency to being transmitted together within a population. If two non-allelic genes, say A and B, with frequencies \( p_A \) and \( p_B \) in the population, are included in the gametes the probability of being transmitted together is

\[
p_{AB} = p_A p_B
\]

with linkage equilibrium

If the genes have linkage disequilibrium, the probability that they are included in the same gamete is more or less than \( p_A p_B \), and linkage disequilibrium is given by

\[
\Delta_{AB} = p_{AB} - p_A p_B
\]

Linkage disequilibrium tends to be dissipated over generations in a random mating population and the rate of dissipation depends on how close the genes are located on the chromosomes or, in other words, on the recombination rate between the two
6.7 Correlation Between Traits and Correlated Response to Selection

Given a recombination rate \( r_1 \) for frequency of recombinants and \( r_0 \) for non-recombinants (where \( r_0 = 1 - r_1 \)), the probability of genes coming together in any generation \( t \) of random mating is

\[
p_{ABt} = r_0 p_{ABt-1} + (1 - r_0) p_{APB}
\]

and the amount of linkage disequilibrium is (Cockerham, 1956)

\[
\Delta_{ABt} = p_{ABt} - p_{APB} = r_0 \Delta_{ABt-1}
\]

Thus the initial amount of linkage disequilibrium is dissipated at the rate of the recombination fraction of each generation of random mating. The approach to linkage equilibrium is gradual and asymptotic; it is faster when the recombination fraction approaches its maximum value (\( r_1 = \frac{1}{2} \)) and it is slower for closely linked genes. When genes are not closely linked, linkage disequilibrium is not an important cause of correlation between characters in random mated populations. In such cases the existence of genetic correlations is mostly attributed to pleiotropic effects.

Environmental correlations also exist because measurements of several traits are taken from the same individual or from the same family. For example, a positive environmental correlation is expected to occur between plant height and ear height in the same plants because a microenvironment that favors plant height also increases ear height and vice versa. When two traits are evaluated by the average (or total) of a family, the environmental deviation in a given plot containing the given family affects all individuals of that plot and causes an environmental correlation of characters among plots.

The linear correlation coefficient \( r \) between variables, say trait 1 and trait 2, is calculated by the following ratio:

\[
r = \frac{\hat{\text{Cov}}(1, 2)}{\hat{\sigma}_1 \hat{\sigma}_2}.
\]

If the two variables are two characters in a maize plant, \( \hat{\text{Cov}}(1, 2) \) is the covariance between the two characters and \( \hat{\sigma}_1 \) and \( \hat{\sigma}_2 \) are their standard deviations. The correlation coefficient is a measure of the degree of association between two characters or the degree to which two characters vary together. It is an estimate of \( \rho \), a parameter of the bivariate normal distribution. The quantity \( r\sqrt{n-2}/\sqrt{1-r^2} \) is distributed as \( t \) with \( (n-2) \) degrees of freedom and is used to test the null hypothesis that \( \rho = 0 \) when \( r \) is calculated from paired values.

In plant breeding, genetic and phenotypic correlations are important. Genetic correlations are related only to genetic causes. The expression genetic correlation is more appropriate when all genetic effects are involved (broad sense) and has a wider use in homozygous self-pollinated and in apomictic species. On the other hand, additive genetic correlation (or simply additive correlation), involving only additive effects, is more appropriate to cross-pollinated species, such as maize, mainly because the information from the correlation is used in connection with recurrent selection. Phenotypic correlation involves both genetic and environmental effects.
Estimation of genetic and phenotypic correlations is based on components of variances and covariances that are estimated from analyses of variance and covariance, respectively. Covariance components are obtained in the same manner as variance components, because for any experimental design the coefficients of expected mean products are the same as those for expected mean squares (Mode and Robinson, 1959). Genetic components of covariance are related to the covariance of relatives in the same way as genetic components of variance and thus can be estimated by the same methods, which are described in Chapter 4. For example, in an analysis of half-sib families the covariance between half-sibs involving two characters is an estimate of one-fourth the additive genetic covariance between them \([\frac{1}{4} \hat{\text{Cov}} A_{12}]\).

An example of estimation of additive genetic variances and covariances is given in Table 6.18, where data and analysis are on an individual plant basis.

**Table 6.18** Analysis of variance and covariance for two tassel traits, number of branches (subscript 1), and weight (subscript 2) for half-sib families in a maize population (Geraldi et al., 1975)

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>MS₁</th>
<th>MP₁₂</th>
<th>MS₂</th>
<th>E(MS)ᵃ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reps/Expt.</td>
<td>4</td>
<td>123.34</td>
<td>10.73</td>
<td>7.35</td>
<td></td>
</tr>
<tr>
<td>Families/Expt.</td>
<td>38</td>
<td>103.74</td>
<td>16.06</td>
<td>15.36</td>
<td></td>
</tr>
<tr>
<td>Pooled error</td>
<td>76</td>
<td>24.57</td>
<td>4.91</td>
<td>5.57</td>
<td></td>
</tr>
<tr>
<td>Withinᵇ</td>
<td>1080</td>
<td>21.36</td>
<td>5.17</td>
<td>4.22</td>
<td></td>
</tr>
</tbody>
</table>

ᵃOr E(MP), by changing variance to covariance
ᵇEstimated from 40 families and \(n = 10\) plants/family

From Table 6.18 the following components of genetic variance and covariance can be estimated:

\[
\hat{\sigma}^2_{A1} = 4\hat{\sigma}^2_{f1} = 4(103.74 - 24.57)/30 = 10.556;
\]

\[
\hat{\sigma}^2_{A2} = 4\hat{\sigma}^2_{f2} = 4(15.36 - 5.57)/30 = 1.305
\]

\[
\hat{\text{Cov}} A_{12} = 4(16.06 - 4.91)/30 = 1.487
\]

Additive genetic correlation is then calculated by

\[
r_A = \frac{1.487/[(1.142)(3.249)]}{0.401}.
\]

Phenotypic variances \(\hat{\sigma}^2_{P1}\) and \(\hat{\sigma}^2_{P2}\) are estimated by \(\hat{\sigma}^2_{f} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{w}\) and the phenotypic covariance by \(\hat{\text{Cov}} P_{12} = \hat{\text{Cov}} f + \hat{\text{Cov}} p + \hat{\text{Cov}} w\); hence the following values are found:
\[ \hat{\sigma}_{y1}^2 = 24.320, \quad \hat{\sigma}_{y2}^2 = 4.681, \quad \text{and} \quad \hat{\text{Cov}}_{y12} = 5.517. \]

Using these estimates the phenotypic correlation coefficient is found to be \( r_p = 0.516. \)

When the experiment includes only two replications, components of genetic variance and covariance can easily be found by using a cross covariance technique as suggested by Falconer and Mackay (1996). Denoting by \( I_1 \) and \( I_2 \) the measurements (half-sib family values in the example given) of traits 1 and 2 in one replication and by \( II_1 \) and \( II_2 \) the measurements of the same traits in another replication, the following relations for half-sib families hold:

\[ \hat{\text{Cov}}(I_1, II_1) = \hat{\sigma}_{f1}^2 = \left( \frac{1}{4} \right) \hat{\sigma}_{A1}^2 \]
\[ \hat{\text{Cov}}(I_2, II_2) = \hat{\sigma}_{f2}^2 = \left( \frac{1}{4} \right) \hat{\sigma}_{A2}^2 \]
\[ \left( \frac{1}{2} \right) \left[ \hat{\text{Cov}}(I_1, II_2) + \hat{\text{Cov}}(I_2, II_1) \right] = \hat{\text{Cov}} f_{12} = \left( \frac{1}{4} \right) \hat{\text{Cov} A_{12}} \]

Additive genetic correlation is important in selection programs because it gives information about the degree of association between two traits by way of additive or breeding values of individuals, which are the effects that can be changed by selection. In other words, selection for one trait will cause a change in its mean through additive effects of genes of selected individuals. If another trait is correlated additively to the first, selection will cause an indirect change in the mean of the second trait. Hence the basic question is as follows:

If selection is for one trait, to what extent will it change the second trait? It depends on the degree of association between the two traits. This indirect change is known as correlated response and can be predicted in the same way as direct response for one trait. The prediction procedure also involves a regression problem. Denoting by \( X_1 \) a measure of trait 1 in the selected individual or family (selection unit) and by \( W_1 \) a measure of the same trait in an individual from the improved population genetically related to \( X_1 \), then for a unit deviation in \( X \) a deviation of \( bW_1 \) (regression coefficient of \( W_1 \) on \( X_1 \)) is expected in \( W_1 \); this regression has been shown to be \( c \hat{\sigma}_{A1}^2 / \hat{\sigma}_{A1}^2 \). Denoting by \( W_2 \) a measure of trait 2 in the same individual in the improved population, then for each unit deviation in breeding value of the selected trait there is an expected deviation of \( bW_2 \) in breeding value of the second trait.

The total change in the second trait or the correlated response is

\[ \text{CR}_2 = \Delta G_1 b_{w2w1} \]
\[ = \left[ s(\hat{\sigma}_{A1}) / \hat{\sigma}_{f1}^2 \right] \left( \hat{\text{Cov} A_{12}} / \hat{\sigma}_{A1}^2 \right) \]
\[ = sc(\hat{\text{Cov} A_{12}}) / \hat{\sigma}_{f1}^2 \]

If truncation selection is used, this formula changes to

\[ \text{CR}_2 = kc \hat{\text{Cov} A_{12}} / \hat{\sigma}_{f1}^2 \]
where $\hat{\text{Cov}} A_{12}$ is the additive covariance between the two traits, $\hat{\sigma}_{P_1}^2$ is the phenotypic variance of the selection unit, and $k$ is a function of the selection intensity.

For example, considering mass selection for one sex, $CR_2$ can be written as

$$CR_2 = k(\frac{1}{2}) \left[ \frac{\hat{\text{Cov}} A_{12}}{(\hat{\sigma}_{A_1} \hat{\sigma}_{A_2})} \right] \left( \frac{\hat{\sigma}_{A_1}}{\hat{\sigma}_{P_1}} \right)$$

$$= (\frac{1}{2}) k r_{A_{12}} h_1 \hat{\sigma}_{A_2}$$

where $h_1$ is the square root of heritability for trait 1.

In the sample given in Table 6.18, the correlated response for tassel weight under mass selection for number of branches at selection intensity of 10% ($k = 1.755$) would be

$$CR_2 = 1.755(\frac{1}{2})(1.487/24.320) = 0.265 \text{ g},$$

which represents a change of 3.6% of the original mean ($m_1 = 7.37 \text{ g}$). For among and within half-sib family selection (at a rate of 1 cycle per year) the expected correlated response would be

$$CR_2 = 1.400(\frac{1}{8})(1.487/1.860) + 1.755(\frac{3}{8})(1.487/4.622) = 0.140 + 0.212 = 0.352 \text{ g}$$

for a selection intensity of 20% among families and 10% within families. This change represents a correlated response in tassel weight of about 4.8% of the original mean.

This example also illustrates another important point relative to correlated traits: the possibility of indirect selection when the primary trait is difficult to measure or evaluate. It has been suggested that selection for smaller tassels would increase plant efficiency in maize (Hunter et al., 1969, 1973; Buren et al., 1974; Mock and Schuetz, 1974), but direct selection for tassel weight would be difficult. The number of branches per tassel can be evaluated more easily and with a relatively high degree of accuracy. The merit of indirect selection relative to direct selection on trait 2 is measured by the ratio of expected correlated response over direct response or, in other words, the relative efficiency of indirect selection compared to direct selection, which is $r_A k_1 h_1/(k_2 h_2)$, where $r_A$ is the additive genetic correlation, $k_1$ and $k_2$ are the selection differentials in standard units, and $h_1$ and $h_2$ are the square roots of the heritability for traits 1 and 2, respectively (Falconer and Mackay, 1996). Indirect selection has an advantage over direct selection when $r_A h_1 > h_2$, i.e., the secondary trait has a higher heritability than the desired character and the additive correlation between them is high. In our example (Table 6.18) indirect selection for tassel weight would not show greater progress than direct selection because $r_A h_1 < h_2$. However, indirect selection should be successful because tassel weight is more difficult to evaluate, additive genetic correlation with number of branches is moderate ($r_A = 0.40$ in our case), and the heritability coefficient for number of branches ($\hat{h}^2 = 0.43$) is substantially higher than for tassel weight ($\hat{h}^2 = 0.28$). Note that heritability
coefficients, as well as additive genetic correlation, depend on the population under selection and on environmental conditions, suggesting that the advantage of indirect selection must be investigated for each particular situation. Information from previous studies relative to heritability, types of gene action, and environmental conditions may serve as a general guide for the breeder’s decisions.

The theory can easily be explained the following way. Selection for a trait \((X)\) can produce correlated responses on another trait \((Y)\) due to the presence of a genetic correlation \((r_A)\) between both traits. The formula for genetic correlation is

\[ r_A = \frac{\hat{\text{Cov} A}}{\hat{\sigma}_{A(X)} \hat{\sigma}_{A(Y)}} \]

\(\text{Cov} A = \text{Additive genetic Covariance between } X \text{ and } Y\)

Genetic correlations can be caused by linkage if loci on the same chromosome control different traits. Random mating, however, can break these linkages. However, if the genetic correlation is due to pleiotropy (same loci controlling traits) the genetic correlation is more stable. Genetic correlations and, therefore, correlated responses can be desirable or not. For instance, selection for high oil content in maize kernels will cause a negative correlated response in kernel starch concentration due to a negative genetic correlation between both traits.

The change of trait \(Y\) based on the selection for trait \(X\) is based on the regression of the breeding value of \(Y\) on the breeding value of \(X\) (change in the breeding value for trait \(Y\) per unit change in the breeding value for trait \(X\)) as follows:

\[ b_{A(YX)} = \frac{\hat{\text{Cov} A}}{\hat{\sigma}_{A(X)}} \]

Hence, replacing \(\hat{\text{Cov} A}\) from the formula above

\[ b_{A(XY)} = r_A \frac{\hat{\sigma}_{A(Y)}}{\hat{\sigma}_{A(X)}} \]

The direct selection response of trait \(X\) can be written as follows

\[ R_X = \frac{k\hat{\sigma}_A^2}{\hat{\sigma}_p^2} \text{ or } R_X = kh_X\hat{\sigma}_{A(X)} \]

Therefore,

\[ \text{CR}_Y = kh_Xh_Yr_A\hat{\sigma}_{p(Y)} \]

So the response of a correlated trait can be predicted if the genetic correlation and the heritability of the two traits are known. As stated before, indirect selection is the selection for a secondary character with the purpose to obtain a positive response in the desirable or primary trait. The relative efficiency of indirect selection compared to direct selection is
In order for correlated response to be successful the correlation among traits should be near 1.0 and the heritability of the second character should be near 0.9. Very few exceptions explain the effectiveness of indirect selection.

6.8 Multi-trait Selection

In practical maize breeding programs, selection for more than one trait is common. The trait of primary importance usually is grain yield, but several other traits such as grain moisture at harvest, lodging resistance, grain quality must be included if the material is to be competitive. Alternatives for the selection of several traits are

(1) Tandem selection
(2) Independent culling levels
(3) Selection indices

Tandem selection emphasizes selection for only one trait for a number of generations. One problem of tandem selection is determining the number of generations of selection to be devoted to each trait. If high selection intensity is used genetic variability decreases rapidly. As a consequence, genetic gain from selection is expected to be reduced after a few generations. Thus the desired limit for each trait, the importance of the trait, and its heritability must be determined before tandem selection is initiated. For example, selection for grain yield would require more generations than selection for plant or ear height. Information about correlated response is also important because a long-term change in one trait may cause an indirect but undesirable change in another important trait. For example, if grain yield and ear height are positively correlated, a long-term selection for grain yield would increase ear height substantially which can be undesirable. An alternative would be to alternate selection, say $M_1$ cycles for trait $X$ followed by $N_1$ cycles for trait $Y$ and again $M_2$ cycles for trait $X$ and so on (Turner and Young, 1969). Tandem selection seems to be useful when the relative importance of each trait changes throughout the years (not the case for most maize traits). If genetic correlations do not exist between say yield and disease and insect resistance, tandem selection can be used effectively to increase the level of disease and insect resistance before selection for yield is initiated.

Independent culling is selection at a given intensity for several traits in the same generation but in sequence for each trait. Suppose that data from 500 half-sib families are to be used as the basis for family selection. Usually these data are family means from replicated trials. The breeder can first select the 200 best families (40%) based on yield. From this sample of 200, a selection intensity of 50% (100 families) can be used for ear height and followed with a 50% selection intensity used for lodging resistance. The total selection intensity would be $0.40 \times 0.50 \times 0.50 = 0.10$ or
10% and only the 50 best families would be used for recombination. This procedure has not been widely used in maize breeding on a formal basis but is often used in applied breeding programs by compromising on different traits across early- and late-generation hybrid trials.

Use of a selection index in its broad sense seems to be common. In most applied selection procedures, breeders’ use an intuitive selection index. By this procedure selection for several traits is made simultaneously and breeders’ decisions are based on the relative weights they give to each trait. Visual observation and experience will improve their decisions. The inherent subjectivity of the selection process enables breeders to put into practice their ability to recognize the desirable genotypes. In this sense plant breeding has been considered an art, especially when making subjective final decisions on releasing desirable genotypes. Genotype evaluation is based on individual plant observation by phenotypic selection and selection efficiency depends on how efficiently breeders apply their empirical weights for several traits. Selection based on family evaluation in replicated trials also requires accuracy in the evaluation of several characters to allow selection to be efficient. Some quantitative characters, such as yield and plant or ear height, can be measured directly but a metrical evaluation is not possible for some traits, such as disease resistance, seedling vigor. For these traits a scale (e.g., from 0 to 5) is commonly used and the accuracy of the evaluation also depends on the breeder’s experience and the amount of compromise made for different traits. If the progeny is exceptionally high yielding, the breeder may select it even though the level of some other trait may not be as high as desired. However, this particular family will continue to be screened per se and as source of inbred line development across years. The family means from replicated trials are further examined. It is also important to observe family performance in each individual replication for all traits: with four replications a given family can have an outstanding performance in three of them but a very poor performance in the fourth, resulting in a relatively low average. This may indicate some abnormality for that family in the fourth replication and previous observation in the experimental field may serve as auxiliary information. Also, consistency of family performance over replications is an important criterion for selection. Another detail has to do with stand variation. If correction is made to compensate for stand variations and the stand in several replications is consistently low, a given family’s mean may be overestimated. In addition, consistently low stand may be an indication of some genetic abnormality. Use of computer programs to list family means and data from individual replications for all traits may facilitate breeders’ work, but decisions about selection still are based on their ability to give to the several traits the appropriate weights they have visualized.

The optimum selection index, first used by Smith (1936) in plants and later by Hazel and Lush (1942) in animals, has not been used extensively in maize breeding. Several authors have reported that use of a selection index would improve selection efficiency relative to selection based on only one trait (Laible and Dirks, 1968; Wolff, 1972; Martin and Salvioli, 1973; Kauffman and Dudley, 1979; St. Martin, 1980). The superiority of the index was reported by Young (1961) when contrasted with tandem selection and independent culling level. It was concluded that the
superiority of the index increases with increasing number of traits under selection but decreases with increasing differences in relative importance. Its superiority was at a maximum when the traits considered were equally important. Gain from selection for any given trait is expected to decrease as additional traits are included in the index, so the choice for traits to be included must be done objectively.

Use of a selection index is most suitable for animals and crop species where the relative value of each character is readily determined through its economic value. In maize, yield is the trait of primary importance, but other traits (e.g., stalk lodging) also have a direct effect on yield when mechanical harvesting is used. Other traits are not measured directly as is yield, but they affect final yield so that assignment of relative weights is a subjective task. Also, precision of variance and covariance estimates are generally low, thus limiting greater use of a selection index. Highly improved populations require better techniques for further improvement and a selection index may be useful when-high-precision parameter estimates that can enhance expected progress are available. The computational procedures for the use of selection indices were originally given by Smith (1936), Hazel and Lush (1942), Hazel (1943), Robinson et al. (1951), Kempthorne (1957), and Brim et al. (1959). The phenotypic value of an individual can be represented by $P = G + E$ for each trait. When considering several traits, it is desirable to choose individuals with the best combination of these traits. The basis for such a selection is the selection index, which takes into account a combination of traits according to their relative weights. Thus each individual has an index value (score) and selection is based on this value. It is possible that the highest yielding individuals will not get the highest scores in the use of a selection index. Because grain yield in maize is the most important trait that has a direct economic value, relative weights for other traits should be given subjectively. Note that even if a weight of zero is given for all traits except yield (weight 1), they will contribute to the optimum index if they are correlated with the primary trait. Such a situation is illustrated by Kempthorne (1957).

The genotypic value of an individual considering several traits is $H_j = a_1 G_{1j} + a_2 G_{2j} + \ldots + a_i G_{ij}$, where $a_i (i = 1, 2, \ldots, n)$ is the relative weight and $G_{ij}$ is the genotypic value of the $j$th individual for the $i$th trait. The objective is to make a selection so that the above combination of characters ($H_j$) is the most desirable. Because the $G_{ij}$ are not known but are evaluated through $P_{ij}$ (phenotypic values), $H_j$ must be evaluated by an index $I_j$ based on phenotypic values such that the correlation between $H_j$ and $I_j$ (denoted by $\rho_{HI}$) is as great as possible. The maximization of $\rho_{HI}$ leads to the following set of equations:

$$GA = PB,$$

where $G$ is a matrix of genetic variances and covariances, $A$ is a vector of a values (relative weight), $P$ is a matrix of phenotypic variances and covariances, and $B$ is a vector of unknown $b$ values. The solution of the set of equations leads to the estimation of $b$ values that give the highest correlation with $H_j$ when applied to the index $I_j$. The $b$ values also maximize the expected progress by selection based on indices (Turner and Young, 1969).
Some modifications have been suggested for use of a selection index. Kempthorne and Nordskog (1959) presented a theory for use of a restricted selection index; i.e., an index constructed such that for \( N = n_1 + n_2 \) traits involved, \( n_1 \) desirable traits are to be changed as efficiently as possible and the \( n_2 \) remaining desirable traits are to be unchanged. Tallis (1962) generalized the theory of a restricted selection index for the case where it is desirable to improve \( n_1 \) traits without limit and \( n_2 \) traits only to a predetermined limit. Williams (1962) suggested the use of a base index that differs from the optimum index because the traits are weighted directly by their economic values. Pesek and Baker (1969) used the concept of “desired genetic gains” to overcome the problem of assigning relative economic weights in crop species. Details for the application of the modified method were presented by Pesek and Baker (1970). A comparison among three different selection indices was reported by Suwantaradon et al. (1975) in maize. Their conclusions were that conventional indices did not give satisfactory improvement for all traits. Base indices were 95% and 97% as efficient as conventional selection indices when relative economic weights were used. Modified selection indices (based on desired gains) were shown to be 46% and 61% as efficient as conventional indices for the specified desired gains and relative economic weights considered.

To address some of the challenges presented additional selection indices were proposed: the multiplicative index (Elston, 1963), the rank-summation index (Mulamba and Mock, 1978), and the heritability (or repeatability for fixed genotypes) index (Smith et al., 1981). The former two are weight free while the heritability index uses the heritability estimates as relative weights and it is easily determined in the analysis of variance. Information on these selection indices and their utilization has been recently updated by Hallauer and Carena (2009).

**References**


Lonnquist, J. H. 1967c. Inter-population improvement: combined S1, mass, and reciprocal recurrent selection. *Maize* 6, CIMMYT.


References


Paterniani, E. 1967a. Inter-population improvement: Reciprocal recurrent selection variations. Maize 8, CIMMYT.


Chapter 7
Selection: Experimental Results

The scientific basis of maize breeding and selection based on planned experiments started at the end of the 19th century. The first methods (mass selection and ear-to-row selection) were soon abandoned because of their supposed lack of efficiency in improving complex traits, such as grain yield. Some other methods and modifications of the previous ones have been used during the 20th century for improvement of populations, mainly after recurrent selection was recognized as a useful method to increase the frequency of favorable alleles in populations (Jenkins, 1940; Hull, 1945; Comstock et al., 1949; Lonnquist, 1949; Sprague and Brimhall, 1950; Jenkins et al., 1954). Also, improvement of populations was recognized as an imperative to increase the probability of obtaining superior and diverse inbred lines and hybrids.

Results of many selection experiments have been reported in the literature and the most important information is generally related to one or more of the following items: (1) effectiveness of selection method for the trait or traits under selection; (2) correlated response of unselected traits; (3) change in genetic variability as measured by additive genetic variance, heritability coefficient, and genetic coefficient of variation estimates; (4) change in combining ability with specific or nonspecific testers; (5) change in inbreeding depression, as evaluated by the performance of inbred progenies; (6) change in heterosis either for intra- or inter-population improvement; (7) comparison of relative effectiveness of different selection methods; (8) effect of environment and genotype–environment interaction; (9) effect of random drift and inbreeding in small populations; (10) value of improved populations as sources of inbred lines for hybrid development; (11) value of environmental control and other techniques (isolation, parental control, etc.) as means for improving selection efficiency; (12) value and efficiency of the selection index; (13) inferences about predominant types of gene action; (14) goodness of fit of quantitative genetic models, e.g., expected progress from selection; and (15) value of different testers for population improvement.
7.1 Measuring Changes from Selection

Observed genetic gain will never be equal to predicted genetic gain due to the importance of genotype by environment interaction and the accuracy of components of variance estimates. In order to make an accurate evaluation of selection response we need to maintain viable seed of each cycle of selection, keep large population samples, approach linkage equilibrium, use effective population sizes when recombinating, and produce seed the year before evaluation.

The objective of most selection experiments is to obtain improved varieties or hybrids; the success itself is a measure of selection effectiveness, which varies with the way by which effectiveness is quantified (Hallauer, 1992; Pandey and Gardner, 1992). As an example, three cycles of full-sib intra-population recurrent selection were conducted in NDSAB(MER)C12, an early-maturing population adapted to the northern USA after 12 cycles of modified ear-to-row selection, to obtain NDSAB(MER-FS)C15. The heritability index was utilized for selecting four traits simultaneously. As a breeder you want to evaluate the progress made and decide to evaluate selection response by measuring changes from selection. Seed of all cycles was produced in 2006 and selection cycles were evaluated in replicated multi-location trials in 2007 and 2008. The average values for grain yield (Mg/ha) were as follows:

\[
\begin{align*}
C0 &= 4.2 \\
C1 &= 4.7 \\
C2 &= 5.9 \\
C3 &= 7.8
\end{align*}
\]

There are different ways to evaluate response to selection:

1. **Total gain method:**

   \[
   R = \frac{(Cn - C0)}{C0} \times 100 = \% \text{Total gain}
   \]

   In this example,

   \[
   R = \frac{(7.8 \text{ Mg/ha} C3 - 4.2 \text{ Mg/ha} C0)}{4.2 \text{ Mg/ha}} \times 100 = 85.7\%
   \]

   However, response to selection can be expressed in gain per cycle or gain per year. Since we have three cycles of selection:

   \[
   R = \% \text{Total gain} / n \quad \text{being "}n\text{" = number of cycles}
   \]

   \[
   R = 85.7\% / 3 = 28.6\%/\text{cycle}
   \]

   The method used for selection needed 2 years per cycle, then:

   \[
   R = \% \text{Total gain} / (n \times y) \quad \text{being "}y\text{" = number of years}
   \]

   \[
   R = 85.7\% / 6 = 14.2\%/\text{year}
   \]
The selection response per year is considered the best estimate for comparing among selection methods when the methods require different numbers of years per cycle.

(2) Regression method:

This is the average response per cycle of selection (regression of desired traits on cycles of selection). The regression coefficient measures the gain per cycle of selection. In the example discussed above the regression coefficient \( b \) is equal to 1.2 Mg/ha/cycle (Fig. 7.1). To obtain the gain in %/cycle, then

\[
R = \frac{b}{C_0}
\]

being “C0” = the mean of the original population

\[
R = \frac{1.2\text{Mg/ha}}{4.2\text{Mg/ha}} = 28.6\% /\text{cycle}
\]

Some researchers use the linear regression coefficient from the performance of populations on the number of selection cycles and the gain is given on a per cycle basis; it can be expressed in the original units, in percent of the observed original mean, or in percent of the original mean predicted by the linear regression. The latter is a good reference point only when there is a good fit of the response to selection and the linear regression model (Fig. 7.1). As seen in the example the gain from selection also can be expressed on a per-year basis by dividing the gain per cycle by the number of years required to perform one cycle of selection. In some instances the gain is expressed in terms of the selection unit, e.g., the gain in performance of \( S_1 \) lines in the \( S_1 \) family selection or the gain in performance in the population × testcross in recurrent selection using a non-related tester. In reciprocal recurrent selection, the most important information is the gain in the crossed population, although changes in the parental populations and in heterosis also are useful if they are not confounded with inbreeding. Another possible measure of the gain from selection is the change in the mean of selection units (families) evaluated in different years and adjusted to a check (constant through cycles) performance. In some instances (usually commercial hybrids) the performance of improved populations or the gain itself is merely expressed in percent of check performance and/or benchmarks. For any unit by which the gain can be expressed, the most precise evaluation of the gain from selection is obtained when the original population, the improved populations representing each cycle, and checks (e.g., commercial hybrids and/or

![Graph showing selection response evaluation of early-maturing maize population NDSAB(MER-FS) through the regression method](attachment:image.png)

\[
y = 1.2x + 2.65 \\
R^2 = 0.94
\]
other reference genotypes) are represented in yield trials replicated over an extensive number of environments. However, for simplicity some preliminary results during the selection program can be reported from the available data, e.g., the performance of family means expressed in percent of checks in each cycle.

The simple or multiple regression models were suggested by Eberhart (1964) to estimate the rate of response for continued selection. For one population undergoing selection, the simplest linear model is

$$Y_i = \mu_0 + \beta_1 x_i + \delta_i$$

where $Y_i$ are observed means over cycles of selection ($i = 0, 1, ..., c$); $\mu_0$ estimates the original population mean; $\beta_1$ is the linear regression coefficient, which expresses the rate of gain per cycle of selection; $X_i$ are cycles of selection; and $\delta_i$ are deviations from the linear model. The least-squares regression analysis, as given by Anderson and Bancroft (1952) or Steel and Torrie (1960), is used for estimation of parameters ($\hat{\mu}_0$ and $\hat{\beta}_1$) and to partition the variation among populations into sums of squares due to linear regression and deviations from the model. If deviations from the model are shown to be non-significant, $\hat{\beta}_1$ provides a good estimation of gain from selection (gain per cycle) and the sum of squares due to $\hat{\beta}_1$, calculated as $R(\hat{\beta}_1) = R(\hat{\mu}_0, \hat{\beta}_1) - R(\hat{\mu}_0)$, provides a test for significance of the observed changes due to selection. If evaluation of effect of quadratic regression is desired, the model is

$$Y_i = \mu_0 + \beta_1 x_i + \beta_2 X_i^2 + \delta_i$$

where $\beta_2$ is the quadratic regression coefficient.

A somewhat different model can be used to fit population means resulting from different selection procedures in the same base population. A linear regression coefficient must be assigned to each selection procedure for estimation of parameters and analysis of variance. A model for fitting two different procedures is $Y_{ij} = \mu_0 + \beta_{11} X_{i1} + \beta_{12} X_{i2} + \delta_{ij}$, or, in general,

$$Y_{ij} = \mu_0 + \sum_j \beta_{1j} X_{ij} + \delta_{ij}$$

where $j = 1, 2, ..., m$ refers to the method of selection. When only one base population is used there is only one parameter $\mu_0$ to express the original population mean. Examples of maize selection programs that fit this model are

(a) divergent recurrent selection in one population and
(b) recurrent selection involving different methods, e.g., $S_1$ and HS family selection.
In the analysis of variance, hypotheses that can be tested are as follows

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Interpretation of hypothesis</th>
<th>Sum of squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0: \beta_{11} = 0$</td>
<td>Ineffective for procedure 1</td>
<td>$R(\beta_{11}) = R(\mu_0, \beta_{11}, \beta_{12}) - R(\mu_0, \beta_{12})$</td>
</tr>
<tr>
<td>$H_0: \beta_{12} = 0$</td>
<td>Ineffective for procedure 2</td>
<td>$R(\beta_{12}) = R(\mu_0, \beta_{11}, \beta_{12}) - R(\mu_0, \beta_{11})$</td>
</tr>
<tr>
<td>$H_0: \beta_{11} = \beta_{12} = \beta$</td>
<td>Equally effective for procedures 1 and 2</td>
<td>$R(\beta_{11} = \beta_{12}) = R(\mu_0, \beta_{11}, \beta_{12}) - R(\mu_0, \beta)$</td>
</tr>
<tr>
<td>$H_0: \beta_{11} = \beta_{12} = \beta'$</td>
<td>Equally effective in opposite directions for procedures 1 and 2</td>
<td>$R(\beta_{11} = -\beta_{12}) = R(\mu_0, \beta_{11}, \beta_{12}) - R(\mu_0, \beta')$</td>
</tr>
</tbody>
</table>

An example is given in Table 7.1, using data from 10 cycles of divergent mass selection for ear length (Cortez-Mendoza and Hallauer, 1979). Deviations from regression were non-significant in the analysis of variance after fitting the linear regression. For testing the four hypotheses, the reduced models were used to solve the normal equations for determining sums of squares for each formulated hypothesis (Table 7.1).

### Table 7.1

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Estimates</th>
<th>Sums of squares</th>
<th>$R$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{11} = 0$</td>
<td>$\mu_0 = 20.91$, $\hat{\beta}_{11} = -0.82$</td>
<td>3925.04, 9.18, 1, 9.18, 42.9</td>
<td></td>
</tr>
<tr>
<td>$\beta_{12} = 0$</td>
<td>$\mu_0 = 16.77$, $\hat{\beta}_{12} = 0.70$</td>
<td>3898.27, 35.95, 1, 35.95, 168.0</td>
<td></td>
</tr>
<tr>
<td>$\beta_{11} = \beta_{12} = \beta$</td>
<td>$\mu_0 = 19.52$, $\hat{\beta} = -0.16$</td>
<td>3834.06, 100.16, 1, 100.16, 468.1</td>
<td></td>
</tr>
<tr>
<td>$\beta_{11} = -\beta_{12} = \beta'$</td>
<td>$\mu_0 = 18.66$, $\hat{\beta}' = 0.48$</td>
<td>3931.42, 2.80, 1, 2.80, 13.1</td>
<td></td>
</tr>
<tr>
<td>Pooled error</td>
<td></td>
<td>436, 0.21</td>
<td></td>
</tr>
</tbody>
</table>

$a\hat{\beta}_{ij}$ refers to $\hat{\beta}_{11}, \hat{\beta}_{12}, \hat{\beta}$, and $\hat{\beta}'$

$bR' = R(\mu_0, \hat{\beta}_{11}, \hat{\beta}_{12}) - R(\mu_0, \hat{\beta}_{11})$

Response to divergent mass selection for increased and decreased ear length was significant, and rates of response also were significantly different for increased and decreased ear length.

More complex models also can be used for inter-population improvement when information about the effects of selection on populations themselves, on the crossed population, and on heterosis in the crossed population is desired. A similar model can be applied when two populations are improved independently by some method
of intra-population recurrent selection. For example, results reported by Paterniani and Vencovsky (1977) on a modified reciprocal recurrent selection program were analyzed on the basis of a more complex model that included the mean of the base populations, heterosis in the original cross, changes in the means of populations themselves, changes in heterosis in crosses between advanced cycles, and the effect of reciprocal crosses. The use of such a model and results are shown in Table 7.2.

Table 7.2  Expected and observed means (yield, t/ha), based on a multiple regression model, of populations (P: Piramex; C: Cateto) and population crosses; and estimates of parameters after reciprocal recurrent selection based on testcross of half-sib families

<table>
<thead>
<tr>
<th>Population</th>
<th>Model for expected meana</th>
<th>Observed</th>
<th>Expectedb</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0</td>
<td>( m_p )</td>
<td>5.641</td>
<td>5.727</td>
</tr>
<tr>
<td>C0</td>
<td>( m_c )</td>
<td>4.109</td>
<td>4.023</td>
</tr>
<tr>
<td>P1</td>
<td>( m_p + g_p )</td>
<td>6.012</td>
<td>5.927</td>
</tr>
<tr>
<td>C1</td>
<td>( m_c + g_c )</td>
<td>4.236</td>
<td>4.323</td>
</tr>
<tr>
<td>P0×C0</td>
<td>( \frac{1}{2}(m_p + m_c) + h + r )</td>
<td>5.834</td>
<td>5.822</td>
</tr>
<tr>
<td>C0×P0</td>
<td>( \frac{1}{2}(m_p + m_c) + h - r )</td>
<td>5.799</td>
<td>5.724</td>
</tr>
<tr>
<td>P1×C0</td>
<td>( \frac{1}{2}(m_p + g_p + m_c) + h + \frac{1}{2}g_h + r )</td>
<td>5.754</td>
<td>6.015</td>
</tr>
<tr>
<td>C1×P0</td>
<td>( \frac{1}{2}(m_p + m_c + g_c) + h + \frac{1}{2}g_h - r )</td>
<td>4.530</td>
<td>5.967</td>
</tr>
<tr>
<td>P1×C1</td>
<td>( \frac{1}{2}(m_p + g_p + m_c + g_c) + h + g_h + r )</td>
<td>6.505</td>
<td>6.258</td>
</tr>
<tr>
<td>C1×P1</td>
<td>( \frac{1}{2}(m_p + g_p + m_c + g_c) + h + g_h - r )</td>
<td>5.998</td>
<td>6.160</td>
</tr>
</tbody>
</table>

Estimates: \( \hat{m}_p = 5.727; \hat{m}_c = 4.023; \hat{h} = 0.898; \hat{r} = 0.049; \hat{g}_p = 0.200; \hat{g}_c = 0.300; \hat{g}_h = 0.185 \)

\( a \) \( g_p, g_c \): expected gain in performance of populations per se, due to selection; \( h \): expected heterosis in original cross; \( g_h \): expected change in heterosis; and \( r \): expected deviation in reciprocal crosses

\( b \) Deviations from model were non-significant; \( R^2 = 0.69 \)

Source: Adapted from Paterniani and Vencovsky (1977)

Least-squares multiple regression procedures have not been used extensively for evaluating selection response. However, as pointed out by Eberhart (1964), availability of general multiple regression computer programs permits the use of very large \( X'X \) matrices and estimation of parameters and analysis of variance can be performed easily.

(3) Model of Smith (1983)

Few attempts have been done to provide direct information for the principles of recurrent selection and to overcome its limitations. One of the goals of recurrent selection within a population is the increase in the frequency of favorable alleles but this frequency has usually not been evaluated directly. Evaluation methods (1) (total gain) and (2) (regression) do not explain how mean changes are related to changes in allelic frequencies and to inbreeding due to selection and genetic drift. The model of Smith evaluates selection response as the average change in allelic frequency. Genetic models based on generation mean analysis have been developed for this purpose (Burton et al., 1971; Hammond and Gardner, 1974; Smith, 1979a, b, 1983).
Gardner and Eberhart (1966) developed a variety diallel model in which means were related with their appropriate genetic expectations. Hammond and Gardner (1974) modified the model in order to evaluate genetic gain after selection. This allowed the estimation of changes in allelic frequencies and the partition of progress from selection into effects due to homozygous and to heterozygous loci. Smith (1979a) proposed a genetic model to interpret changes in the population mean under selection not only as a function of allelic frequencies and allelic effects but also based on the importance of inbreeding depression as a limit to potential gains. The model requires data from selection cycles, selfed cycles, crosses and selfed crosses of two populations and has the power to obtain estimates of the contributions of additive as well as dominance effects to population means, estimates of heterosis, and estimates of realized response to selection adjusted for genetic drift. Basic assumptions of the model include Hardy–Weinberg equilibrium, diploid inheritance, and no epistasis. The first application of the Smith’s model was in the reciprocal recurrent selection (RRS) among BSSS and BSCB1 populations. Eberhart et al. (1973) reported non-significant regression coefficients for the evaluation of BSSS(R) and BSCB1(R) populations per se. The model, however, showed significant values for the average changes in allelic frequencies, which indicated the effectiveness of RRS in increasing the frequency of favorable alleles. Considerable inbreeding depression was reported as being responsible for the non-significant changes in their means during selection (Smith, 1979a). Half-sib selection in BSSS showed significantly less inbreeding depression due to genetic drift than BSSS using RRS although the same number of lines was used for recombination for both methods (Smith, 1979a). Inbred progeny selection also has exhibited less inbreeding depression than RRS (Oyervides-García and Hallauer, 1986). This can be explained because selection for grain yield based on S_2 progenies produced important changes in frequency of additive effects while the changes in frequency of dominance alleles were more important after eight cycles of reciprocal recurrent selection. Tanner and Smith (1987) found, however, that inbred progeny selection was associated with higher values of genetic drift while Garay et al. (1996) did not obtain a significant estimate of genetic drift after two cycles of S_1 recurrent selection. Stojšin and Kannenberg (1994) evaluated five populations and four different selection methods and reported that the type of change that occurs in allelic frequencies was determined more by the selection method than by population type. This gives some insight to the differential influence of selection methods on type of changes. Smith (1979b) evaluated half-sib and S_1 recurrent selection in BSK (data from Burton et al., 1971) with his proposed model. Mean changes were found to be a function of inbreeding depression, and they concluded that the effects of small population sizes may be larger than the effects due to selection. Genetic drift and selection are the two main forces affecting allelic frequencies, and, therefore, the effectiveness of any selection method will depend on their balance (Stojšin and Kannenberg, 1994). Estimates of genetic drift under the Smith’s model have been reported to be negative for yield in most studies (Iglesias and Hallauer, 1989; Eyherabide and Hallauer, 1991b; Keeratinijakal and Lamkey, 1993; Stojšin and Kannenberg, 1994). Evaluation of selection response using the Smith’s model has provided more information on the causes of grain yield
Table 7.3  Genetic gains of grain yield in maize populations based on estimates from the Smith’s model (Smith, 1979a, b, 1983)

<table>
<thead>
<tr>
<th>Population</th>
<th>Adjusted&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Unadjusted</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS(R)</td>
<td>0.356</td>
<td>0.344</td>
<td>Helms et al. (1989)</td>
</tr>
<tr>
<td>BSSS(R)</td>
<td>0.259</td>
<td>0.118</td>
<td>Oyervides-G and Hallauer (1986)</td>
</tr>
<tr>
<td>BSSS(R)</td>
<td>0.296</td>
<td>0.284</td>
<td>Keeratinijakal and Lamkey (1993)</td>
</tr>
<tr>
<td>BSSS(R)</td>
<td>0.135</td>
<td>−1.170</td>
<td>Smith (1979a)</td>
</tr>
<tr>
<td>BSSS(HT)</td>
<td>0.344</td>
<td>0.331</td>
<td>Helms et al. (1989)</td>
</tr>
<tr>
<td>BSCB1(R)</td>
<td>0.297</td>
<td>0.285</td>
<td>Keeratinijakal and Lamkey (1993)</td>
</tr>
<tr>
<td>BSCB1(R)</td>
<td>1.480</td>
<td>−0.920</td>
<td>Smith (1979a)</td>
</tr>
<tr>
<td>BSCB1(R)</td>
<td>0.372</td>
<td>0.336</td>
<td>Smith (1983)</td>
</tr>
<tr>
<td>BS13(HT)</td>
<td>0.496</td>
<td>0.483</td>
<td>Smith (1983)</td>
</tr>
<tr>
<td>BS13(S)</td>
<td>0.226</td>
<td>0.207</td>
<td>Helms et al. (1989)</td>
</tr>
<tr>
<td>BSK(S)</td>
<td>0.468</td>
<td>0.335</td>
<td>Tanner and Smith (1987)</td>
</tr>
<tr>
<td>BSK(S)</td>
<td>0.197</td>
<td>0.014</td>
<td>Smith (1979b)</td>
</tr>
<tr>
<td>BSK(HT)</td>
<td>0.122</td>
<td>−0.060</td>
<td>Smith (1979b)</td>
</tr>
<tr>
<td>BS10(FR)</td>
<td>0.360</td>
<td>0.350</td>
<td>Eyherabide and Hallauer (1991b)</td>
</tr>
<tr>
<td>BS11(FR)</td>
<td>0.220</td>
<td>0.210</td>
<td>Eyherabide and Hallauer (1991b)</td>
</tr>
<tr>
<td>EZS1(S)</td>
<td>0.439</td>
<td>0.394</td>
<td>Garay et al. (1996)</td>
</tr>
<tr>
<td>EZS2(S)</td>
<td>0.588</td>
<td>0.568</td>
<td>Garay et al. (1996)</td>
</tr>
<tr>
<td>CGSynA(R)</td>
<td>0.120</td>
<td>0.117</td>
<td>Stoššin and Kannenberg (1994)</td>
</tr>
<tr>
<td>CGSynB(S)</td>
<td>0.080</td>
<td>0.080</td>
<td>Stoššin and Kannenberg (1994)</td>
</tr>
<tr>
<td>CGG(HS)</td>
<td>0.080</td>
<td>0.070</td>
<td>Stoššin and Kannenberg (1994)</td>
</tr>
<tr>
<td>CGN(S)</td>
<td>0.220</td>
<td>0.218</td>
<td>Stoššin and Kannenberg (1994)</td>
</tr>
<tr>
<td>CGW(S)</td>
<td>1.540</td>
<td>1.330</td>
<td>Stoššin and Kannenberg (1994)</td>
</tr>
<tr>
<td>BSSS(R)</td>
<td>0.276</td>
<td>0.254</td>
<td>Smith (1983)</td>
</tr>
<tr>
<td>Leaming(S)</td>
<td>0.352</td>
<td>0.275</td>
<td>Carena and Hallauer (2001)</td>
</tr>
<tr>
<td>Midland(S)</td>
<td>0.184</td>
<td>0.120</td>
<td>Carena and Hallauer (2001)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Values adjusted for genetic drift

improvement. A comparison of genetic gain estimates after evaluating populations per se is shown in Table 7.3.

In most selection experiments in which prediction of gain was possible, predicted values tended to be greater than observed gains. Probably the most important factors are sampling error of predictors and genotype by environment interaction effects. Prediction of gain from selection is based on the heritability coefficient, which is a ratio of variances. It is also well known that variance estimates are generally associated with large errors, which are reflected in the error of predicted gains. One way to reduce sampling error would be to average the estimates of several cycles. The procedure introduces a bias because the genetic situation is expected to change from cycle to cycle, but if the bias is small relative to the reduction in sampling error the average may be closer to the true value than are individual predictions (Moll, 1959). Genotype by environment interaction causes an upward bias in the predicted gain, as already shown in Chapter 6. Other possible causes of discrepancies may be
random drift and inbreeding in small populations, linkage disequilibrium, and epistasis. Therefore, adding more genetic parameters in models increases our predictive power (Smith, 1983; Dudley and Johnson, 2009). In spite of the fact that predicted gains are generally greater than observed gains, a good agreement through cycles of selection has been observed in some instances.

7.2 Improvement from Intra-population Selection

Recurrent selection studies in maize were developed to obtain improved sources of germplasm for the potential extraction of lines and their possible use in hybrids (Hallauer, 1985). The development of improved inbred lines depends on the improvement of sources of germplasm, both genetically narrow- and broad-based populations. An important similarity is that both recurrent selection and inbred line development in maize use evaluation strategies. Both breeding methods need to be complementary and they will only be successful if both are given the same importance.

7.2.1 Mass (or Individual) Selection (M)

Plant breeding programs emphasize the genetic improvement of grain yield. Even though there have been cases where mass selection has improved yield and other genetically complex traits, mass selection has shown to be extremely effective in modifying highly heritable traits that are easy to measure. Among breeding methods, mass selection is probably the best for resource allocation as it has a minimum cost, it allows maximum sample sizes for the lowest selection intensity, it provides maximum additive genetic variance, and a genetic progress of 1 year per selection cycle. The ineffectiveness of mass selection often claimed for maize population improvement was not because of the method per se but due to poor isolation, lack of environmental control, and poor plot techniques. Therefore, proper isolation techniques, sample sizes, and experimental precision are as essential as the trait being selected. Therefore, mass selection is definitively an effective method to improve traits under minimum resources.

Probably, the most successful example for mass selection genetic improvement across populations is the one obtained with flowering time in maize. Even so expensive efforts to identify and isolate genes (e.g., \textit{vtgI}) to use in MAS have been proposed. Moreover, Buckler et al. (2009) recently evaluated “almost a million” genotypes across eight environments and they concluded this trait is very complex genetically with numerous QTL having mostly additive gene action. However, even though the genetic control of this trait seems to be complex at the genotypic level, flowering time is probably one of the simplest traits to measure and the one that has shown most genetic progress with mass selection. The choice of genetic materials (e.g., NAM population) might not represent a large variation for the trait. In addition, this trait had 48 estimates gathered across the literature confirming the importance of additive genetic variance (Table 5.1). Stratified mass selection (Gardner,
based only on the female parent (e.g., parental control = 0.5) has been a cost-effective-proven methodology for adapting exotic germplasm into breeding programs and target environments. Tropical populations ETO Composite (BS16), Antigua Composite (BS27), Tuxpeno Composite (BS28), and Suwan-1 (BS29) were effectively adapted to temperate environments (Table 7.4). Similar results ($b = -3.0$ days/year) were obtained for each of the four tropical varieties. Within each variety average days to silking decreased 3 days per cycle of selection and within 6–8 cycles of selection the four tropical varieties were adapted to central Iowa (Hallauer, 1999). Similar unpublished results were achieved with Tuson Composite introduced from Cuba.

Temperate elite populations BS11(FR)C13 [NDBS11(FR-M)C3], BSK(HI)C11 [NDBSK(HI-M)C3], and the cross BS10(FR)C13 × BS11(FR)C13 (NDBS1011) were also effectively adapted to early-maturing temperate environments (Carena et al., 2008). Again, similar results across 3 years ($b = -2.0$ days/year) were obtained for each of the three elite varieties. Within 3–4 cycles of selection the three improved populations were adapted to North Dakota conditions (e.g., <95 RM) (Table 7.5). Similar unpublished results were obtained with a composite of four highland tropical maize improved populations [NDSHLC(M)C4].

Correlated changes with selection for earlier silk emergence were, in general, positive. In Iowa experiments they included reduced plant and ear heights, reduced tassel and leaf sizes, root and stalk lodging, and lower incidence of *Ustilago maydis*; grain yield significantly increased because of better adaptation to temperate environments. In North Dakota experiments (Eno and Carena, 2008) correlated changes with selection for earlier silk emergence included significant reduction in plant and ear heights, days to pollen shed, grain moisture at harvest, as well as significant increases in grain yield and test weight across populations because of better adaptation to US northern temperate environments. Correlated responses were, in many cases, similar to those obtained in the Iowa experiments (e.g., plant and ear heights). However, yield components varied according to the genetic background. For instance, cob diameter was reduced for NDBSK(HI-M)C3, the adapted version of BSK(HI)C11; and cob diameter and kernel-row number were reduced for NDBS11(FR-M)C3, the adapted version of BS11(FR)C13. However, prolificacy was reduced while ear length, ear diameter, cob diameter, and kernel-row number were significantly increased in NDBS1011, the adapted version of BS10(FR)C13 × BS11(FR)C13. Besides, the improvement for earliness in BS11(FR)C13 has reduced its original high oil content.

Sample sizes for each cycle of selection ranged from 15,000 (Iowa selection studies) to 22,500 (North Dakota selection studies) evaluating more than a million plants at a very low cost. Effective population sizes ranged from 250 to 400 individuals. Complex genetic mechanisms may be involved (Buckler et al., 2009), but it certainly seems that major genes had a large role in changing flowering time based on selection for early silk emergence.

Mass selection for earlier flowering is not new. Troyer and Brown (1972) selected three late semi-exotic synthetics (adapted approximately to 39° latitude) for early flowering. In a further report (Troyer and Brown, 1976) selection also was effective
Table 7.4  Response to mass selection for earlier silking in ETO Composite, Antigua Composite, Tuxpeno Composite, and Suwan-1 maize cultivars and correlated responses for ear height and grain yield

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Days to silk no.</th>
<th>Ear height cm</th>
<th>Days to silk no.</th>
<th>Ear height cm</th>
<th>Grain yield q/ha</th>
<th>Days to silk no.</th>
<th>Ear height cm</th>
<th>Grain yield q/ha</th>
<th>Days to silk no.</th>
<th>Ear height cm</th>
<th>Grain yield q/ha</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>116</td>
<td>212</td>
<td>91</td>
<td>143</td>
<td>7.0</td>
<td>95</td>
<td>131</td>
<td>43.3</td>
<td>105</td>
<td>141</td>
<td>41.0</td>
</tr>
<tr>
<td>C1</td>
<td>112</td>
<td>192</td>
<td>91</td>
<td>146</td>
<td>12.5</td>
<td>90</td>
<td>101</td>
<td>56.0</td>
<td>99</td>
<td>131</td>
<td>56.3</td>
</tr>
<tr>
<td>C2</td>
<td>110</td>
<td>182</td>
<td>82</td>
<td>137</td>
<td>37.2</td>
<td>86</td>
<td>93</td>
<td>58.0</td>
<td>96</td>
<td>124</td>
<td>61.7</td>
</tr>
<tr>
<td>C3</td>
<td>106</td>
<td>178</td>
<td>79</td>
<td>133</td>
<td>46.1</td>
<td>81</td>
<td>86</td>
<td>56.0</td>
<td>93</td>
<td>120</td>
<td>54.5</td>
</tr>
<tr>
<td>C4</td>
<td>100</td>
<td>146</td>
<td>76</td>
<td>121</td>
<td>50.9</td>
<td>79</td>
<td>81</td>
<td>50.0</td>
<td>90</td>
<td>114</td>
<td>67.8</td>
</tr>
<tr>
<td>C5</td>
<td>—</td>
<td>—</td>
<td>74</td>
<td>117</td>
<td>50.4</td>
<td>79</td>
<td>81</td>
<td>58.0</td>
<td>92</td>
<td>121</td>
<td>62.0</td>
</tr>
<tr>
<td>C6</td>
<td>—</td>
<td>—</td>
<td>74</td>
<td>124</td>
<td>50.9</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Bc</td>
<td>−3.8</td>
<td>−15</td>
<td>−3.2</td>
<td>−5</td>
<td>19.3</td>
<td>−3.3</td>
<td>−9</td>
<td>3.0</td>
<td>−2.6</td>
<td>−4</td>
<td>5.9</td>
</tr>
</tbody>
</table>

aBS16, BS27, BS28, and BS29 are the new pedigree names for adapted versions
bNumber of days from planting to 50% silk emergence
cEstimates of linear response over cycles of mass selection (at a rate of one cycle per year)
Source: Adapted from Hallauer and Carena (2009)
### Table 7.5

Response to mass selection for earlier silking in BSK(HI)C11, BS11(FR)C13, and BS10(FR)C13 × BS11(FR)C13 maize cultivars and correlated responses for grain moisture at harvest and grain yield across nine North Dakota environments

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Days to silk</th>
<th>Grain moisture</th>
<th>Grain yield</th>
<th>Days to silk</th>
<th>Grain moisture</th>
<th>Grain yield</th>
<th>Days to silk</th>
<th>Grain moisture</th>
<th>Grain yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>72.0</td>
<td>28.1</td>
<td>55.0</td>
<td>73.5</td>
<td>28.0</td>
<td>56.8</td>
<td>77.2</td>
<td>32.4</td>
<td>45.0</td>
</tr>
<tr>
<td>C1</td>
<td>69.6</td>
<td>25.9</td>
<td>54.0</td>
<td>72.5</td>
<td>26.7</td>
<td>58.0</td>
<td>73.1</td>
<td>28.7</td>
<td>57.0</td>
</tr>
<tr>
<td>C2</td>
<td>69.0</td>
<td>24.0</td>
<td>59.0</td>
<td>70.3</td>
<td>24.0</td>
<td>59.0</td>
<td>71.6</td>
<td>26.4</td>
<td>59.0</td>
</tr>
<tr>
<td>C3</td>
<td>65.1</td>
<td>21.9</td>
<td>66.0</td>
<td>67.9</td>
<td>20.8</td>
<td>66.0</td>
<td>70.2</td>
<td>24.0</td>
<td>63.0</td>
</tr>
<tr>
<td>C4</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>79</td>
<td>—</td>
<td>—</td>
<td>69.1</td>
<td>21.5</td>
<td>60.0</td>
</tr>
<tr>
<td>Bc</td>
<td>−2.3</td>
<td>−2.1</td>
<td>3.7</td>
<td>−3.3</td>
<td>−2.4</td>
<td>3.0</td>
<td>−2.6</td>
<td>−2.7</td>
<td>3.8</td>
</tr>
</tbody>
</table>

| a | NDBSK(HI-M)C3, NDBS11(FR-M)C3, and NDBS1011 are the new pedigree names for adapted versions |
| b | Number of days from planting to 50% silk emergence |
| c | Estimates of linear response over cycles of mass selection (at a rate of one cycle per year) |

Source: Adapted from Eno and Carena (2008)

### Table 7.6

Results of selection for early flowering expressed as the gain on average basis (a) or by the linear regression coefficient (b) in percent of the original population for several traits

<table>
<thead>
<tr>
<th>Latitude of adaptation</th>
<th>Days to silk</th>
<th>Grain moisture</th>
<th>Plant height</th>
<th>Ear height</th>
<th>Silk delay</th>
<th>Yield</th>
<th>Broken stalks</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I) 39°a</td>
<td>(a) −2.3</td>
<td>−3.2</td>
<td>−2.1</td>
<td>−5.2</td>
<td>−7.6</td>
<td>1.9</td>
<td>47.2</td>
</tr>
<tr>
<td></td>
<td>(b) −2.4</td>
<td>−3.3</td>
<td>−1.9</td>
<td>−3.2</td>
<td>−7.1</td>
<td>2.2</td>
<td>47.0</td>
</tr>
<tr>
<td>(II) 41.5°</td>
<td>(a) −2.3</td>
<td>−4.4</td>
<td>−1.9</td>
<td>−3.0</td>
<td>−12.4</td>
<td>1.8</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>(b) −2.2</td>
<td>−4.2</td>
<td>−1.8</td>
<td>−2.8</td>
<td>−11.4</td>
<td>2.0</td>
<td>31.4</td>
</tr>
<tr>
<td>(III) 44°</td>
<td>(a) −1.1</td>
<td>−4.0</td>
<td>−1.4</td>
<td>−2.6</td>
<td>−7.3</td>
<td>−3.2</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>(b) −1.1</td>
<td>−3.7</td>
<td>−1.5</td>
<td>−2.8</td>
<td>−6.8</td>
<td>−3.3</td>
<td>10.0</td>
</tr>
</tbody>
</table>

| a | (I) Three late semi-exotic synthetics adapted to 39°; (II) seven southern Iowa adapted synthetics to 41.5°; and (III) three synthetics adapted to 44° latitude |
| b | Source: Adapted from Troyer and Brown (1972, 1976); Troyer (1975) |

Flowering time in maize can also be improved through inbred line development methods derived from large F₂ populations from early × late crosses (Rinke and Sentz, 1961; Carena et al., 2009) or even larger F₂ populations derived from late × late crosses (e.g., the NDSU maize breeding program has developed early-maturing lines from late × late crosses). Troyer (1976) also selected the earliest 2% of plants to flower in 18 F₂ populations. In one group of eight populations (whose parents had
similar flowering dates), the effect of selection was evaluated after sib-mating two cycles and selfing the third cycle and then tested in crosses with Iowa inbred B14. Selection effect per cycle averaged 4% selection index increase, 340 kg/ha yield increase, 0.6% grain moisture decrease, and 0.6 days less to flower. The second group (10 F2 populations from five elite lines), selected in the same manner, was evaluated in crosses with all possible (15) double crosses among the lines; selection effect per cycle averaged 250 kg/ha yield increase, 1% grain moisture decrease, 3.7% stalk breakage increase, 7.0 cm plant height decrease, 1.2 days less to flower, and 0.1 day less silk delay. Obilana (1974) detected no improvement in selecting for earliness when evaluation was over three environments. However, when evaluation was in only one location, days to tassel were reduced from 52 to 46, a gain of 11.5% in two cycles. The third cycle gave an additional gain of 5.8% in the same location. No significant increase was observed in the number of ears per plant, plant and ear height, and grain yield.

Smith (1909) first reported results from divergent selection for ear placement in the Leaming open-pollinated variety of maize. Selection for high and low ear placement was in adjacent blocks planted ear-to-row where selection for ear height was only within the best yielding rows. Results for two traits are summarized in Table 7.7.

<table>
<thead>
<tr>
<th>Year</th>
<th>Ear height (cm)</th>
<th>Yield (q/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>1903</td>
<td>143.2</td>
<td>108.7</td>
</tr>
<tr>
<td>1904</td>
<td>127.8</td>
<td>97.3</td>
</tr>
<tr>
<td>1905</td>
<td>160.8</td>
<td>105.7</td>
</tr>
<tr>
<td>1906</td>
<td>143.8</td>
<td>64.8</td>
</tr>
<tr>
<td>1907</td>
<td>183.9</td>
<td>84.3</td>
</tr>
<tr>
<td>1908</td>
<td>145.5</td>
<td>58.7</td>
</tr>
</tbody>
</table>

Because the data were obtained in different years, the most adequate measure of effect of selection on ear height is provided by the difference between high and low selection blocks. The difference increased at a rate of 15.7 cm per generation, as measured by the linear regression coefficient. Under the hypothesis that selection for high and low ear height was equally effective, the change would be about 7.9 cm per generation in each direction. The effect of selection on yield cannot be measured because high yield was taken into account in selection for low and high ear height. Nevertheless, 3 years of data indicated no great differences between the two strains for divergent selection for ear height.

Mass selection for ear height has shown a good rate of progress in some selection programs. Vera and Crane (1970) reported a rate of gain of about 4.5% per cycle after two cycles of selection for low ear height in two populations. Subsequently, Acosta and Crane (1972) observed a reduction of about 24% in ear height in both
selected populations after four cycles at selection intensity of 20%. Josephson and Kincer (1973) reported a reduction in ear height of 18.3 cm in a late synthetic and 25.9 cm in an early synthetic after 10 generations of selection for lower ear height. Subsequently, Josephson et al. (1976) reported that the reduction in ear height was 3.18 cm per cycle in both synthetics after 12 generations of selection for lower ear height. Changes in other traits also were observed in these reports. In each case reduction in ear height was followed by reduction in plant height and yield. A positive association between yield and ear height or plant height was also observed when selection was for yield by Gardner (1969a), Hallauer and Sears (1969), and Darrah et al. (1972). However, as seen before, yield might not be associated with plant and ear height when adapting germplasm the same way grain moisture at harvest does not seem to be associated with yield when conducting divergent mass selection for ear moisture (Cross et al., 1987). In a similar way, Cross (1990) was able to select individuals for rapid leaf expansion rate based on a single environment without delaying maturity.

Kiesselbach (1922) reported results of several selection experiments. Selection for high, medium, and low ear height was conducted in contiguous blocks for 5 years and resulted in average ear heights of 137, 119, and 112 cm, respectively, with very small differences in yield (32.4, 33.3, and 33.6 q/ha). Similarly, selection for low and high lodging resulted in 22 and 29% (on the average) of lodged plants, respectively. Selection for three levels of erectness of ears (erect, medium erect, and dropping ears) resulted in 42, 29, and 21% erect ears as averaged over the period. During 6 years of selection for ear type relative to height, diameter, and texture (smooth or rough) there was not a clear trend of selection when comparisons were made for yield. Two years of selection for deep-grained, rough ears and shallow-grained, smooth ears were compared for yield with the original Hogue’s Yellow Dent open-pollinated variety; on the average the grain yields were, respectively, 40.2, 43.7, and 40.5 q/ha. Additional data of selection for long, smooth ears of Reid Yellow Dent open-pollinated variety contrasted for yield with standard, medium rough ears of the same variety showed that the smooth type surpassed the rough type by 2.5 q/ha. In all experiments for ear type for 12 different years the long, smooth ears surpassed all others in yield. However, the author concluded from limited evidence that ear-type considerations are rather neutral in their relation to yield. However, they seem to have certain importance when separating heterotic groups.

Lonnquist (1952) and Lonnquist and McGill (1956) reported the effect of visual selection for desirable plants in advancing four synthetic varieties of maize. They concluded that a slight improvement in yield was observed for all synthetics and that gain in yield was accompanied by slightly later maturity. It is quite possible that in advancing synthetic varieties or variety composites many breeders around the world have observed some gain in yield and other traits by simple or visual mass selection; e.g., adaptive mass selection in semi-exotic populations resulted in an increase in yield of about 5% per year accompanied by an increase in prolificacy (Mathema, 1971). Also, effective selection for early silking was reported by Hallauer and Sears (1972) in a program for integrating exotic germplasm in the Corn Belt maize program, the same way seen across mass selection studies for early flowering time.
Paterniani (1974b) observed very little increase in yield (2.3%) in advancing a cross between two populations by two generations of mass selection. Thus mass selection seems to be very useful in adapting exotic or semi-exotic populations or in advancing generations of synthetic varieties or composites. Mild selection for important quantitative traits like lodging resistance, disease resistance, prolificacy, ear height, and other plant and ear characters, as well as for some qualitative traits, will surely result in some improvement in the population and possibly some increase in yield. Therefore, mass selection is the most suitable method in these cases to effect some improvement without drastic loss of genetic variability because of the large effective population size that can be maintained. Populations obtained by this method are usable as base populations in selection experiments (including mass selection) that use higher selection pressures. Applied maize breeding programs often have adaptation of elite germplasm as one of their long-term program goals before maximizing their genetic improvement with other selection methods. Therefore, mass selection can be useful for exotic populations to breeding programs before initiating recurrent selection programs based on progeny testing.

Mass selection for yield in maize was more effective after the modification introduced by Gardner (1961). After four cycles of mass selection in the Hays Golden open-pollinated variety (control and irradiated), using the grid system to minimize the effect of environment, an average gain per year of 3.9% over the parental control variety was obtained resulting in total gain of about 15%. Lonnquist (1961) reported that an additional cycle in the same variety showed evidence of continued improvement, increasing total gain to about 19%. Subsequent reports of the same experiment (mass selection in Hays Golden variety) indicated additional gain after continued selection. Thus, Lonnquist et al. (1966) reported a 12.7% improvement for yield after six cycles or a rate of 2.1% per cycle. Gardner (1968) observed a gain of 2.7% per cycle after 10 generations. In later reports Gardner (1969a, 1973) showed a gain of 3.0% per cycle after 15 generations, as predicted by the linear regression over cycles. This was in excellent agreement with the expected 3.1% response. Gardner (1976) reported results from 19 generations in the control and irradiated Hays Golden variety. Both selected populations tended to plateau at about generation 13 (Fig. 7.2).

The trend of response did not depart significantly from a linear regression. At generation 17 yields started to decrease and the population never recovered the yields attained at cycle 15 in subsequent cycles of selection. Gardner (1976) discussed some possible causes of the sudden decrease in generation 17, summarized as follows. The original population may have been contaminated or unintentionally selected during its periodical (one in each 4 years) multiplication, but there was no evidence that the population had changed. Another hypothesis is that the shift from isolated nurseries with uniform soil to extremely variable soils and other environmental factors may have precluded an accurate identification of superior genotypes. Severe heat and drought in the last cycles of selection accentuated that problem. Also, previous selection could lead to favorably linked epistatic combinations of genes for which selection was no longer effective in extreme environmental variations, leading to reduced yields. Still another hypothesis is related to genotype
by environment interaction in the later years of testing. It was observed during the experiment that Hays Golden has the capacity to perform well under less favorable conditions, but plasticity of response is not found in the selected material. Thus the selected generations would interact more with environment than the original Hays Golden variety, leading to a relative decrease in yield of the selected material. Undesirable correlated responses due to selection for only one trait could be another problem. However, it is interesting to note that a similar trend was obtained after 10 cycles of ear-to-row selection method in the same population (see Fig. 7.3).

Johnson (1963) obtained an increase of 33% after three cycles in a tropical variety, an average of 11% per cycle. Also, Eberhart et al. (1967) increased yield of Kitale Composite A from 51.2 to 55.0 q/ha, a gain of 7.4% in one cycle of selection. Vencovsky et al. (1970) reported a progress of 3.8 and 1.7% per cycle in two Brazilian populations, Paulista Dent (five cycles) and Cateto M. Gerais (three cycles), respectively. Other experiments have shown encouraging results from mass selection for yield components. As seen in previous chapters, genotypic and phenotypic correlations between yield and its components tend to be large between yield and its components. These genetic parameters can be improved on certain selection environments. For instance, increasing plant density can increase the association between yield and prolificacy even though the selection environment is not favorable to high expression of prolificacy. Prolificacy (more than one ear per plant) has been proposed for indirect selection of yield and other traits due to its advantages in improving yield, stability (e.g., more stalk sugar concentration during flowering), anthesis-silk interval, earliness, height, lodging, seedling vigor, nitrogen use efficiency, and even photosynthetic efficiency. Prolificacy has also been instrumental for many recurrent selection programs.

Arboleda-Rivera and Compton (1974) developed improved populations through mass selection for three different seasonal conditions. Selection in the rainy season increased grain yield and prolificacy at rates of 10.5 and 8.8% per cycle, respectively, when tested in rainy seasons (direct response). However, when tested

Fig. 7.2  Response to mass selection for grain yield in Hays Golden (Gardner, 1976)
in dry seasons (indirect response) the increases were only 0.8 and 1.0% per cycle, respectively. Selection in dry seasons resulted in a direct response for yield of 2.5% per cycle, which increased to 7.6% per cycle when tested in rainy seasons; similarly, the gain in ears per plant was 4.4% and 11.4%, respectively. In the population tested under both seasons, rainy and dry, gains for yield were 5.3 and 1.1% per cycle, respectively, and for prolificacy the respective gains were 7.0% and 3.3% per cycle. The effects of 12 cycles of mass selection for prolificacy in the open-pollinated variety Golden Glow were given by Coors and Mardones (1989). Direct response for number of ears per plant was significant with average number of ears per plant increasing from 0.94 for the C0 to 1.25 for the C12. Correlated changes with selection for increased ear number included significant increases for grain yield (0.29 t/ha/cycle), performance index, and plant height. The correlated increases in grain yield after selection for prolificacy are in agreement with Lonnquist (1967), Mareck and Gardner (1979), Singh et al. (1986), Subandi (1990), most of them utilizing mass selection. However, only one out of four selection programs for prolificacy showed yield increases in Carena et al. (1998) with the exception that inbred progeny selection was utilized instead. Coors and Mardones (1989) concluded that with proper techniques, mass selection is an effective method for germplasm improvement programs. A similar experiment was conducted by Rodriguez et al. (1976). Mass selection in population MB-21 was performed in three ways: selection in the less favorable season (MB-21-A); selection in the more favorable season (MB-21-B); and alternate selection in both seasons (MB-21-AB). The average gains of yield per cycle were 3.3% (two cycles), −1.0% (two cycles), and 4.9% (four cycles), respectively. The greater increase in prolificacy (1.8% per cycle) also was observed in the population selected in both seasons (MB-21-AB), whereas decreases in prolificacy of 0.4 and 3.2% per cycle were obtained from selections in MB-21-A and MB-21-B, respectively. It was inferred that less favorable environmental conditions allowed greater responses to selection which agrees with the findings of Carena et al. (1998). Lonnquist (1967) showed that selection for prolificacy in the Hays Golden variety gave an increase in the number of ears per plant and in grain yield; a gain of 6.3% per cycle in yield was observed after five cycles of selection at 5% selection intensity, so total gain is comparable with that reported by Gardner (1969b) after 10 cycles of mass selection at 10% selection intensity in the same variety. Torregroza and Harpstead (1967) also reported effectiveness of divergent mass selection for prolificacy; selection for increasing prolificacy resulted in 14% greater yield and 28% more ears per plant, while selection for single ear plants resulted in 5% decrease in yield and 7% decrease in ears per plant. Results from a similar experiment reported by Torregroza (1973) after 11 generations of selection showed 48% increase in prolificacy and 35% increase in yield in selection for high prolificacy; on the other hand, decreases of 16% in prolificacy and 7% in yield resulted from selection for single ear plants. Torregroza et al. (1976) evaluated advanced inter-varietal crosses MB-51 and MB-56 after nine and four cycles of mass selection for prolificacy, respectively. The average response was an increase from 1.3 (original) to 1.6 (ninth cycle) ears per plant in MB-56. Gains per cycle (regression coefficients) were $3.4 \pm 0.3\%$ and $2.0 \pm 0.8\%$ for MB-51 and MB-56, respectively; the corresponding
Selection: Experimental Results

Gains per cycle for yield were 5.5 ± 1.2% and 3.0 ± 1.7%, respectively. Kincer and Josephson (1976) reported a 13.2% increase in ears per plant and 33.1% increase in yield after five generations of selection for prolificacy in a program started after nine generations of selection for yield in the variety Jellicorse. Eleven cycles of mass selection for increased prolificacy was completed in the BS11 population (Weyhrich et al., 1998). Mass selection was effective for increased ear number per plant, but compared with six other methods of recurrent selection, mass selection had the lowest response (0.6% per cycle) for grain yield improvement when compared with half-sib (1.6% per cycle), reciprocal full-sib (2.6% per cycle), and S2 progeny (4.5% per cycle) recurrent selection methods.

Selection for ear length was concluded to be ineffective by Williams and Welton (1915); they stated, ‘It appears that within a variety, at least, the length of an ear of corn is largely a matter of environment and cannot be expected to influence materially succeeding generations.’ This conclusion was qualified as unjustified by Sprague (1966), because ear length is a highly heritable trait (see Table 5.1). Selection was effective in separating the original population Iowa Long Ear Synthetic into short ear and long ear subpopulations (Hallauer, 1968). After 10 generations, Cortez-Mendoza and Hallauer (1979) observed an increase of 0.32 cm (1.6%) and a decrease of 0.64 cm (3.2%) per year in the long ear and short ear subpopulations, respectively; the asymmetry of response was probably due to the higher frequencies of genes for long ear length in the original population, which was synthesized from long ear inbred lines. After 27 cycles of selection, Hallauer et al. (2004) reported continued response to longer ($b = 0.27 ± 0.03$ cm) and shorter ($b = −0.37 ± 0.03$ cm) ear length with correlated decreases in grain yield with selection for longer ($b = −0.01 ± 0.01$ t/ha) and shorter ($b = −0.08 ± 0.01$ t/ha) ear lengths. Other correlated responses with selection for longer ears included significant reductions for ear diameter, kernel-row number, and kernel depth and significant increases for plant and ear height and days to flower; opposite correlated changes were reported for the same traits with selection for shorter ear length.

Other yield components were targeted for mass selection. Moreira et al. (2008) found that 20 years of mass selection for ear size (level of fasciation) in the Portuguese variety Pigarro was effective for ear size but the selection methods used were not sufficient to significantly increase grain yield. Padgett et al. (1968) obtained about 2% gain per cycle in weight per 100 kernels in each direction after nine cycles of divergent mass selection for seed size. A similar lack of response for yield was observed when selecting for seed size (Odhiambo and Compton, 1987). Cross and Djava (1985) conducted four cycles of mass selection for kernel depth in early-maturing population NDSAB with the goal of improving yield. Selection was successful to increase kernel depth, kernel weight, and yield but it also produced an undesirable correlated response on test weight.

Genter (1976b) reported that after 10 cycles of mass selection in a composite of Mexican races, yield increased 171%, with a gain per cycle (regression coefficient) of 19.1%. Increase in yield was accompanied by a decrease in plant height, ear height, days to silk, grain moisture, smutted plants, and days from pollen shed to silk emergence; no change was detected for root lodging but stalk lodging increased.
Josephson and Kincer (1976) reported results of 14 generations of mass selection for yield in the variety Jellicorse; no increase was observed in the first 4 generations, maximum increase per generation was 13.1%, and no increase was shown beyond the 10th generation. Similar increases in yield were obtained by Osuna-Ayalla (1976) in two populations after six cycles of selection. Estimated gains per cycle (linear regression coefficient) were 2.8 and 3.5% for Dent Composite and Flint Composite, respectively. Other short-term selection experiments have demonstrated the effectiveness of mass selection for yield. Miranda et al. (1971) reported results of selection for high yield (two cycles) and for low yield (one cycle) in selection either for one sex or for both sexes in the population Dent Composite. Two cycles of selection for high yield were not effective (1.7% increase) when based only on the female parent (one sex) but were effective (7% increase) when based on both sexes. One cycle of selection for low yield, however, was not effective when selecting for both sexes but was highly effective (15.7% decrease) when selecting for only one sex. In spite of discrepancies in results, a combined least-squares analysis showed that selection for only male gametes (effect of detasseling the poorest plants) was effective as well as was selection for only female gametes (no detasseling). The limited number of cycles precludes any general conclusion. Hakim et al. (1969) obtained an average progress of 4% in only one cycle of selection in the Philippines; when the selected material was evaluated in the same season the gain was 9% over the original. Also, one cycle of mass selection resulted in significant increases by two methods of mass selection. Ordinary stratified mass selection resulted in progress of 11.6%. In the second procedure a constant genotype represented by a double-cross hybrid was planted in alternate hills of the white opaque-2 variety (the population under selection). At harvest the yield of each opaque-2 plant was compared with the yield of the adjacent double-cross hybrid plant, and a progress of 5.6% resulted by this selection procedure. Theoretical considerations suggest that a greater efficiency of the second method relative to stratified mass selection is attained whenever variation among plants of the constant genotype is smaller than environmental variation among strata. A full advantage would be attained by using a single cross as the constant genotype. Simple mass selection in an opaque-2 population conducted by Palma and Burbano (1976) resulted in an increase of 3.3 and 8.2% in the first and second cycles, respectively.

A less pronounced response to mass selection for grain yield has been reported by some authors. Hallauer and Wright (1967) obtained an increase of 4.5% after three cycles for the open-pollinated variety Iowa Ideal; an increasing trend in grain moisture, root lodging, and ear droppage was also observed. In a subsequent report (Hallauer and Sears, 1969), small progress was obtained after selection in Krug (six cycles) and Iowa Ideal (five cycles) with a non-significant increasing rate of about 1.5% per cycle. Darrah et al. (1972, 1978) obtained rates of progress of 0.9 and 0.38 q/ha (0.8%) at 2% selection intensity after 2 and 6 years of mass selection, respectively, in the population Kitale Composite A. At a 10% selection intensity the rates of progress were 0.8 and 0.53 q/ha (1.1%) for 2 and 6 years of selection, respectively. In a population of Nigerian Composite B, progress of 16% was
obtained by Obilana (1974) after four cycles of selection for yield. Long-term mass selection for grain yield and stalk lodging resistance in NDSM(M), NDSAB(M), and NDSCD(M) early-maturing synthetic varieties was not successful, since there were no significant changes in grain yield or stalk lodging in these populations at either low or high densities (Hyrkas and Carena, 2005). In this case, even though selection was performed at low plant densities no genotype by plant density interaction was observed. A summary of results obtained by three or more cycles of mass selection for grain yield is presented in Table 7.8.

Selection for several other traits has been reported in the literature. Ariyanayagan et al. (1974) observed a change in leaf angle of about 3.8° and 10.2° per generation of selection (measured by two methods of leaf angle determination) in a

<table>
<thead>
<tr>
<th>Population</th>
<th>Selection intensity (%)</th>
<th>Number of cycles</th>
<th>Average gain/cycle (%)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hays Golden</td>
<td>10.0</td>
<td>15</td>
<td>3.0</td>
<td>Gardner (1976)</td>
</tr>
<tr>
<td>Tropical</td>
<td>4.7</td>
<td>3</td>
<td>10.3</td>
<td>Johnson (1963)</td>
</tr>
<tr>
<td>Paulista Dent</td>
<td>20.0</td>
<td>5</td>
<td>3.8</td>
<td>Vencovsky et al. (1970)</td>
</tr>
<tr>
<td>Cateto M. Gerais</td>
<td>20.0</td>
<td>3</td>
<td>1.7</td>
<td>Vencovsky et al. (1970)</td>
</tr>
<tr>
<td>Jellicorse</td>
<td>—</td>
<td>14</td>
<td>0.9</td>
<td>Josephson and Kincer (1973)</td>
</tr>
<tr>
<td>Dent Composite</td>
<td>10.0</td>
<td>6</td>
<td>2.8</td>
<td>Osuna-Ayalla (1976)</td>
</tr>
<tr>
<td>Flint Composite</td>
<td>10.0</td>
<td>6</td>
<td>3.4</td>
<td>Osuna-Ayalla (1976)</td>
</tr>
<tr>
<td>Krug</td>
<td>7.5</td>
<td>6</td>
<td>1.6</td>
<td>Hallauer and Sears (1969)</td>
</tr>
<tr>
<td>Iowa Ideal</td>
<td>7.5</td>
<td>5</td>
<td>1.4</td>
<td>Hallauer and Sears (1969)</td>
</tr>
<tr>
<td>Kitale Composite A</td>
<td>2.0</td>
<td>6</td>
<td>0.8</td>
<td>Darrah et al. (1978)</td>
</tr>
<tr>
<td>Kitale Composite A</td>
<td>10.0</td>
<td>6</td>
<td>1.1</td>
<td>Darrah et al. (1978)</td>
</tr>
<tr>
<td>Nigerian Composite B</td>
<td>10.0</td>
<td>4</td>
<td>1.9</td>
<td>Obilana (1974)</td>
</tr>
<tr>
<td>Composite of Mexican races</td>
<td>6.0</td>
<td>10</td>
<td>19.1</td>
<td>Genter (1976b)</td>
</tr>
<tr>
<td>Mezcla Varietales Amarillosa</td>
<td>5.0</td>
<td>3</td>
<td>2.5 (7.6)</td>
<td>Arboleda-Rivera and Compton (1974)</td>
</tr>
<tr>
<td>Mezcla Varietales Amarillosb</td>
<td>5.0</td>
<td>3</td>
<td>0.8 (10.5)</td>
<td>Arboleda-Rivera and Compton (1974)</td>
</tr>
<tr>
<td>Mezcla Varietales Amarillosc</td>
<td>5.0</td>
<td>3</td>
<td>1.1 (5.3)</td>
<td>Arboleda-Rivera and Compton (1974)</td>
</tr>
<tr>
<td>MB-21a</td>
<td>5.0</td>
<td>2</td>
<td>3.3</td>
<td>Rodriguez et al. (1976)</td>
</tr>
<tr>
<td>MB-21b</td>
<td>5.0</td>
<td>2</td>
<td>−1.0</td>
<td>Rodriguez et al. (1976)</td>
</tr>
<tr>
<td>MB-21c</td>
<td>5.0</td>
<td>4</td>
<td>4.9</td>
<td>Rodriguez et al. (1976)</td>
</tr>
<tr>
<td>NDSM</td>
<td>1.0</td>
<td>7</td>
<td>2.3</td>
<td>Hyrkas and Carena (2005)</td>
</tr>
<tr>
<td>NDSAB</td>
<td>1.0</td>
<td>17</td>
<td>−0.2 (0.8)d</td>
<td>Hyrkas and Carena (2005)</td>
</tr>
<tr>
<td>NDSCD</td>
<td>1.0</td>
<td>10</td>
<td>0.2 (2.2)</td>
<td>Hyrkas and Carena (2005)</td>
</tr>
</tbody>
</table>

aSelected for dry, rainy, and both seasons, respectively; the respective gains refer to test in dry season or in wet season (in parentheses)
bSelection for less favorable, more favorable, and both seasons, respectively; gains refer to average test over both seasons in 1 year
cSelection for less favorable, more favorable, and both seasons, respectively; gains refer to average test over both seasons in 1 year
dGain within parentheses is for selection stopping at cycle 14 (NDSAB) and at cycle 7 (NDSCD)
bidirectional phenotypic selection for this trait; types selected toward erect leaves were shorter in plant height, later in maturity, greater in lodging resistance, and greater in light transmissibility. Grain yield variations attributable to leaf angle differences were small and non-significant. Zuber et al. (1971) reported that the percentage of ears with earworm damage was reduced 2.8% per generation after 10 generations of selection for earworm (Heliothis zea Boddie) resistance in two synthetics. Dudley and Alexander (1969) observed changes in yield and plant and ear heights after selection for good seed set and agronomic type in four synthetics. Hanson (1971) selected for high and low stalk volume by use of phenotypic selection within full-sib families of the open-pollinated variety Jarvis. After six cycles the ratio between high vs. low stalk volume was 1.42 for fresh weight and 1.32 for dry weight. It was concluded that selection modified stalk and leaf production primarily and root production secondarily. Net photosynthetic rates were not associated with differential productivity supporting the hypothesis that net photosynthetic rate did not play a major role in determining differential productivity within the adapted population. Subsequently Hanson (1973) reported the effect of selection for stalk volume on the efficiency and number of chloroplasts. Selection for high stalk volume resulted in less chlorophyll per gram of fresh weight but more total chlorophyll than selection for low volume. An identical pattern was found for the measure of DNA. Apparently selection resulted in similar chloroplast number per cell but selection for high stalk volume produced more chloroplasts per leaf. No apparent differences in physical characteristics of chloroplasts were found between selections.

The procedure used by Sprague and Brimhall (1950) and Sprague et al. (1952) for selection of oil content in maize kernels is a type of mass selection with recombination of remnant selfed seed. This procedure follows: A given number of plants were selfed and the seed was analyzed for oil content. A sample of ears with the highest oil percentage was grown ear-to-row and all possible inter-crosses were made by hand, providing seed for the next cycle of selection. This was called the recurrent series. For comparison, a sample of the selfed ears was used to continue selfing and selection (selfing series). Results reported by Sprague et al. (1952) in the population Iowa Stiff Stalk Synthetic show that the recurrent series was more effective in increasing oil content in the kernel than the selfing series. In the recurrent series the change was from 5.0 to 7.0% or 0.4% per year, while in the selfing series the mean oil percentage increased from 5.0 to 5.6 during five generations, an average of 0.13% oil per year. The procedure used by Jenkins et al. (1954) did not include selfing selected plants. However, pollination was controlled because a sample of pollen of selected plants was collected and mixed in apparently equal proportions and used to pollinate the same selected plants, resulting in selection for both sexes. Results after three cycles of selection demonstrated that this recurrent selection method was effective to increase the frequency of alleles for resistance to northern leaf blight, Helminthosporium turcicum Pass.

Smith and Brunson (1925) concluded that mass selection was as effective as ear-to-row selection in improving yield over 10 cycles of selection in the Reid
Yellow Dent variety. The average performances over 10 years were 106.8 and 109.3% of the original variety for the strains maintained by mass selection and by ear-to-row selection for high yield, respectively. In either case the gain was very small.

Mass selection is the oldest method used in plant selection, and the method continues to play an important role in maize breeding. Mass selection has shown to be effective in pre-breeding, at least for adaptation and initial improvement of populations. It seems, however, that after initial increases in the frequency of major alleles response to selection either decreases or reaches a plateau (Gardner, 1976). Therefore, mass selection should be complemented by other progeny selection methods to maximize genetic improvement.

7.2.2 Half-Sib Family Selection (HS)

Long-term continuous selection is essential for germplasm improvement. However, choice of germplasm for long-term genetic improvement might limit the success of germplasm enhancement programs. Half-sib family selection has been the most widely used selection method in genetic improvement. Except for mass selection, a form of half-sib selection (ear-to-row selection) has been used longer than the other methods. The Illinois oil and protein study is the longest selection experiment in maize, and Dudley and Lambert (2004) have provided an interesting summary of the selection study.

7.2.2.1 Ear-to-Row Selection (ER)

Use of half-sib families as progeny tests for maize improvement was introduced by Hopkins (1899) in the Illinois program for quality improvement of the maize kernel at University of Illinois. The open-pollinated variety Burr’s White was used as the base population and the goal of the experiment was to determine if the chemical composition of the maize kernel could be altered by selection. The program started with 163 open-pollinated ears analyzed for oil and protein content from which four strains were selected: high protein, high oil, low protein, and low oil (Hopkins, 1899). Each strain was grown in a separate isolated field and selected according to ear-to-row procedure during nine generations. Approximately, 20% of the ears analyzed were selected. From generations 10 to 25, alternate rows in each breeding plot were detasseled and 20 ears for analysis were taken from the six highest yielding rows. Four ears per row were saved. In 1921 the system was again altered because yielding ability was disregarded and two seed ears were chosen from each of 12 detasseled rows. After generation 28, a system of intra-strain reciprocal crossing between sub-strains was introduced and natural pollination was replaced by controlled hand pollination (Leng, 1962). This breeding procedure continued 100 generations and selection for oil and protein content was effective during the course of this long-term experiment (Dudley and Lambert, 2004). The first report of selection results was presented by Hopkins after two cycles of selection. Subsequently,
periodical reports showed continued progress from selection in the four strains (Smith, 1909; Winter, 1929; Woodworth et al. 1952; Leng 1962; Dudley et al., 1974; Dudley and Lambert, 2004). Dudley et al. (1974) showed that after 70 generations of selection the mean oil content of the high oil and low oil strains were 16.6 and 0.4%, which are 354.8 and 8.5%, respectively, of the oil content in the original variety. Mean protein content in the high protein and low protein strains was 26.6 and 4.4% representing, respectively, 244.0 and 40.4% of the original variety protein content.

After generation 37, a system of random mating or non-selection was compared with the regular breeding system for eight generations (1934–1941). During this period neither continued selection nor the relaxation of selection in the high protein strain produced any significant changes. In the low protein strain, selection produced significant progress for low protein content but no change was observed by random mating. The same pattern was observed in the high oil strain, i.e., significant change by selection but no change by random mating during the first four generations. After four generations a dramatic increase in oil content occurred in both strains. The selected population returned to the previous level in the following year, but the non-selected population remained relatively high in the subsequent years (Leng, 1962).

After generation 48, reverse selection was initiated in all strains in parallel with the normal selections. At generation 70, a 5.6% change was observed in the reverse high oil strain (8.9%) but only 1.6% was observed in the reverse low oil strain (2.4%) as compared with the oil content before reverse selection of 13.5 and 0.7%, respectively. The reverse high protein strain had 8.5% and the reverse low protein strain had 9.6% protein after 22 cycles of selection, as compared to 19.2 and 5.1% protein, respectively, before reverse selection began. After seven generations of reverse selection, the reverse high oil strain was again subdivided, including selection toward high oil (switchback selection). At generation 70, the mean oil content had increased to 14.0%. Dudley and Lambert (1969) reported good agreement between the average expected and the actual gain in the four strains. Results are summarized in Table 7.9.

Dudley (1976) reported that after 76 generations, gains from selection for high oil and low oil strains were 279 and 92% of the original population mean. Similarly, in the high protein and low protein strains, gains of 133 and 78%, respectively, were achieved. Dudley also analyzed several other genetic aspects of the Illinois oil

| Table 7.9 Predicted and observed response per generation from selection for oil and protein content over 65 generations (Dudley and Lambert, 1969) |
|-------------------------------|----------------|---------------|----------------|----------------|
| Direction | Oil content (%) | Protein content (%) |
| High | Predicted | Observed | Predicted | Observed |
|       | 0.13 | 0.16 | 0.26 | 0.22 |
| Low  | −0.02 | −0.07 | −0.16 | −0.09 |
and protein selection experiments; he concluded that enough genetic variation was present to expect future response to selection.

Dudley and Lambert (2004) summarized a detailed analysis of the 100 generations of selection within the high and low strains for oil and protein within the open-pollinated variety Burr’s White (Hopkins, 1899). In their analyses, they partitioned the 100 generations into five segments based on the generations of selection. Significant responses per generation were realized in both percentage of oil and protein in most instances for all segments but the responses after segment 1 tended to be smaller. Change per generation was 0.22 ± 0.07 for high oil for segment 1 vs. 0.16 ± 0.01 for segment 5, and change per generation was 0.30 ± 0.14 for high protein for segment 1 vs. 0.10 ± 0.06 for segment 5. Realized heritability estimates and number of effective genetic factors also decreased with selection for higher percentages of oil and protein. Their evidence suggested that an upper limit had not been attained for higher percentage of oil, but no significant increase was obtained in percentage of protein since generation 88. Total gain in the high oil strain (17.1%) was about four times the total gain for the low oil strain (4.2%). Total gain for the high protein strain (19.6%) was three times the gain for the low protein strain (6.3%). The differences in total gain were not surprising given that zero would be the lower limit for low oil and protein. The successful improvement of grain oil and grain protein is explained partially by the identification of at least 40 QTL for both traits with small effects which agrees with earlier estimates based on classical quantitative genetic theory (Dudley et al., 2007). On the other hand, the NAM population found less than half of the QTL which might indicate challenges in how to sample and choose populations for certain traits and unique alleles.

Some results from use of half-sib family selection (ear-to-row) were reported in the early part of this century. The first experiments were encouraging, but further disappointing results caused the method to be considered ineffective in improving grain yield in maize. Smith (1909) used the ear-to-row system to select for yield among rows and high and low ear heights within rows. Therefore, the first criterion was selection for yield among rows and the second the height of ear for selection (phenotypic) within rows. The effect of selection on yield seemed to be non-significant as indicated by data from 3 years of testing (see Table 7.7). Some recent results have shown that selection for low ear height usually is followed by a decrease in yield in adapted germplasm (Hallauer et al. 2004). In Smith’s experiment, selection for low ear height was especially effective. Hence it is plausible to assume that primary selection for yield was compensated for by any downward change caused by selection for low ear height, although no precise comparison is possible from the data. Smith also selected for erect and drooping ears, using the same procedure as for ear height selection. He concluded that after 5 years the trait had changed to a considerable extent, while no apparent change was observed in yield.

Williams (1907) reported the results of several ear-to-row tests for grain yield. He also pointed out the advantage of using remnant seed from selected rows instead of selecting directly in the ear-to-row test. By use of remnant seeds it was shown that the Ohio Standard Leaming open-pollinated variety gave a yield of 4.5 q/ha.
in excess of the original Leaming. On the other hand, selection within the highest yielding rows of the Clarage open-pollinated variety gave an excess of 1.8 q/ha over the original stock. In both instances comparisons were made in adjacent blocks without replications; consequently, their precision was very poor. Hartley (1909) reported that ears selected from high-yield breeding rows yielded 11.3 q/ha, or 16%, in excess over yields from a general field of maize that was planted in alternate rows to the selected ears from the previous year. On the other hand, Noll (1916) observed that seeds from high-yielding ear rows produced less than seeds from the general field in a selection experiment with College White Cap open-pollinated variety. In a similar experiment with the 90-Day Clarage variety, both sources produced practically equivalent yields. In later experiments involving 90-Day Clarage, seeds from different rows were planted separately. In a test in 1913, the greatest gain was 2.9 q/ha, while two out of six plots yielded less than field-grown seed. In 1915 all plots from high-yielding rows outyielded field-grown seed, and in one instance the excess was 8.0 q/ha. In 1914 remnant seed of the best ears was planted and crossed. Five ears from each cross were planted ear-to-row in the variety field. In both instances the block of field-grown seed used as the check was superior in yield, showing no progress because of selection.

Kiesselbach (1922) reported results of the ear-to-row system at the Nebraska Experiment Station. The following procedures were used to increase high-yielding strains:

1. Continuous ear-to-row selection within the most productive ear-to-row strains.
2. Increasing highest yielding original ears by using remnant seeds in isolated blocks to avoid contamination.
3. Mixing several productive ear-to-row strains and increasing thereafter in a single isolation block.
4. Natural crossing of high yielding ear-to-row strains.

Ear-to-row selection in Hogue’s Yellow Dent did not show significant differences in a comparative test over 7 years (1911–1917). Average yields were 33.4 q/ha for continuous selection, 29.9 q/ha for a single high-yielding strain increased by using remnant seeds, 34.5 q/ha for a composite of four strains, 34.2 q/ha for the intercrossing of four strains, and 33.6 q/ha for the original Hogue’s Yellow Dent. Selection in the Nebraska White Prize variety, using the same procedures, averaged 39.7, 38.1, and 40.7 q/ha using procedures (1), (3), and (4), respectively; the original population yielded 40.0 q/ha.

Smith and Brunson (1925) used the ear-to-row procedures in a divergent selection program for high and low yield in a Reid Yellow Dent variety. The same population was also maintained by mass selection in an isolated block. After 10 years, the high-yielding strain (selected via ear-to-row) averaged 39.1 q/ha and the low-yielding strain 32.5 q/ha; these values represent 109.3 and 90.9% of the original variety. Mass selection resulted in a 6.8% increase over the original variety in the same period. It was concluded that ear-to-row selection for low yield was more effective than selection for high yield. However, precision of the comparisons was very poor, which was common in early maize experiments.
In general, lack of adequate field plot techniques for selection and for comparison of results from selection, associated with other factors such as inbreeding (due to small populations) and lack of isolation, caused the ear-to-row method to be regarded as powerless for improvement of maize yield. Only after modifications introduced by Lonnquist (1964) the method was again regarded as very promising for population improvement. Similar to the modifications proposed by Gardner (1961) on mass selection, Lonnquist (1964) intended to get better estimates of environmental effects.

7.2.2.2 Modified Ear-to-Row Selection (MER)

The modified ear-to-row method (Lonnquist, 1964) is based on among and within half-sib family selection. The first results were published by Paterniani (1967). After three selection cycles within the population Paulista Dent, the improved material yielded 42% in excess over the original population with a regression coefficient of 13.6% over selection cycles. Webel and Lonnquist (1967) reported a gain of 9.4% per cycle relative to the parental variety Hays Golden after four cycles of selection. The first cycle was obtained from selection among and within full-sib families, since the tested progenies resulted from crosses between individual plants within the population; half-sib selection started with open-pollinated ears obtained from the recombination block of full-sib families. In the second cycle, data from the crossing (recombination) block were used as additional information for among-family selection. In subsequent cycles the crossing block was not the source for family performance data, thus providing better opportunities for selection within families because individual plants could be evaluated and selected for grain yield.

Other reports also have shown that the modified ear-to-row procedure is an effective method for the improvement of maize populations. Eberhart et al. (1967) reported results of two cycles of ear-to-row (selection only among families) in three populations; selection was effective for Kitale II (2.8% per cycle) and for Ec573 (11.4% per cycle) but not for Kitale Composite A-Synthetic 1. Paterniani (1969) reported a gain per cycle of about 3.8% (linear regression coefficient) in the population Piramex. Darrah et al. (1972) reported gains varying from 0.9 to 3.2 q/ha per year in four populations (KII, Ec573, H611, and KCA). Darrah et al. (1978) observed different trends in response to selection in the populations KII, Ec573, and H611: 0.83, 2.59, and −0.43 q/ha per year, respectively, after 6 years of selection. The population Kitale Composite A (KCA) gave a response of 1.9 q/ha per year (5.2%) in 10 years as measured by the average response in four selection experiments (Darrah, 1975). Hakim et al. (1969) obtained 6% improvement in one cycle of selection in the Philippines.

Troyer et al. (1965) reported effectiveness of ear-to-row selection for adaptation of an introduced population; over 6 years they observed an increase in yield, a decrease in grain moisture, and a decrease in stalk and root lodging. Paterniani (1974a) obtained an increase in yield of about 5% after two cycles of mass selection and one cycle of among and within half-sib family selection in the population ESALQ-HV1. Relatively small progress was obtained by Lima et al. (1974) after
two cycles of selection in two populations; progress of about 3 and 2% was obtained for the Flint Composite and Dent Composite, respectively, where selection was also based on other traits (ear height and stalk lodging). Paterniani (1974a) obtained a substantial gain (about 35%) after one cycle of selection in Piranão, an open-pollinated brachytic population. Sevilla (1975) evaluated eight cycles of selection in the variety PMC-561, obtaining an average gain of 9.5% per cycle. Segovia (1976) reported a 3.2% gain per cycle after three cycles of selection in the variety Centralmex, but no additional gain was detected in the following three cycles.

Selection in the Hays Golden population, first reported by Webel and Lonnquist (1967), was continued for 10 generations and results were reported by Compton and Bahadur (1977). An observed response (linear regression estimate) of 5.3% per generation was in good agreement with the expected 4.9% response. Gardner (1976) reported results up to 12 generations and a curvilinear response over all cycles was observed, as shown in Fig. 7.3.

The linear regression estimate of the gain per cycle was 4.6%, which is in good agreement with the predicted response of 4.5%, but the trend of response was clearly not linear with a better fit to a quadratic response curve. The relative decrease in yield in the last generations was not expected. Decrease in additive genetic variance is not a sufficient explanation because it could cause the population to level but not to decrease in yield. A change in selection criterion in generation eight by using a selection index based on lodged plants and dropped ears might account for lack of continued response. Additional data showed an average gain of 1.8% per cycle in the population IAC-1 (Miranda et al., 1977). Long-term modified ear-to-row selection was effective for grain yield improvement in NDSAB. Grain yield increased non-linearly at a rate of 2.5% per cycle. Vasal et al. (1982) improved five gene pools at

![Fig. 7.3 Response to half-sib family (modified ear-to-row) selection in Hays Golden (Gardner, 1976)](image)
CIMMYT after four to nine cycles of modified ear-to-row selection of grain yield, flowering time, and ear height. Average response per cycle across the five gene pools was 479.8 kg/ha (12.5%) for grain yield, 2.4 days (3.1%) earlier flowering, and 13.4 cm (12.7%) shorter ear height relative to the original (C0) gene pools. A summary of results from some half-sib selection experiments for yield based on the modified ear-to-row procedure in Table 7.10.

**Table 7.10** Effect of half-sib family (modified ear-to-row) selection (three or more cycles) on yield in several populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Selection intensity (%)</th>
<th>Number of cycles</th>
<th>Gain per cycle (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paulista Dent</td>
<td>15.0</td>
<td>3</td>
<td>13.6</td>
<td>Paterniani (1967)</td>
</tr>
<tr>
<td>Piramex</td>
<td>23.7</td>
<td>4</td>
<td>3.8</td>
<td>Paterniani (1969)</td>
</tr>
<tr>
<td>Hays Golden</td>
<td>20.0</td>
<td>4</td>
<td>9.4</td>
<td>Webel and Lonnquist (1967)</td>
</tr>
<tr>
<td>Hays Golden</td>
<td>20.0</td>
<td>10</td>
<td>5.3</td>
<td>Compton and Bahadur (1977)</td>
</tr>
<tr>
<td>Hays Golden</td>
<td>20.0</td>
<td>12</td>
<td>4.6</td>
<td>Gardner (1976)</td>
</tr>
<tr>
<td>Kitale Composite A</td>
<td>—</td>
<td>6</td>
<td>2.2</td>
<td>Darrah (1975)</td>
</tr>
<tr>
<td>Centramex</td>
<td>22.5</td>
<td>3</td>
<td>3.2</td>
<td>Segovia (1976)</td>
</tr>
<tr>
<td>IAC-1</td>
<td>15.5</td>
<td>7</td>
<td>1.8</td>
<td>Miranda et al. (1977)</td>
</tr>
<tr>
<td>NDSAB</td>
<td>33.3</td>
<td>12</td>
<td>2.5</td>
<td>Hyrkas and Carena (2005)</td>
</tr>
</tbody>
</table>

*aRefers to subsequent reports of the same selection program.

### 7.2.2.3 Half-Sib Family Selection with the Use of Testers (HT)

Selection based on testcross evaluation with a tester (e.g., industry line, populations) is also a type of half-sib family selection. The testcrosses represent half-sib families, which are evaluated in replicated trials across locations. Testers are utilized to develop the half-sib families and the genetic base of testers was the original distinction of the original suggestions of recurrent selection with the use of half-sib families. Jenkins (1940) proposed the source population as tester as a method to improve general combining ability (GCA) whereas Hull (1945) suggested the use of either an inbred or a single cross as tester to improve specific combining ability (SCA). Also, testers could be related or unrelated (Hallauer, 1975b). The latter is the typical procedure from applied breeding programs.

Iowa Stiff Stalk Synthetic (BSSS) and B73 are diverse products that were developed in the public sector before intellectual property and research foundations were common at land-grant universities. B73 definitively was worth its investment on its long-term breeding program at Iowa State University. Most important B73 is the most successful example of integrating pre-breeding with cultivar development (Carena, 2008). However, it may not be the best example for sequencing unique alleles not present in the B73 genome (Carena et al., 2009). Half-sib family selection for GCA was initiated in BSSS with double-cross hybrid IA13 as the tester, the
population was named BS13 (Hallauer, 1992). Half-sib family selection was continued in BS13 until 1970 when the method of selection was changed to inbred progeny recurrent selection based on $S_1$–$S_2$ early generation lines. The half-sib family phase of recurrent selection in BS13 was effective in identifying inbred lines B14, B37, B73, and B84, which have been widely used in hybrids and as germplasm in pedigree breeding to develop recycled inbred lines (Mikel and Dudley, 2006).

Lonnquist (1949) selected $S_1$ lines based on their performance as testcrosses with the parental variety Krug. The testcrosses represented half-sib families, which were evaluated in replicated trials, and remnant $S_1$ seeds of the male plants from the best testcrosses were recombined. Although the half-sib families are obtained from the cross $S_1$ progenies $\times$ populations, they are expected to have the same array of genotypes as $S_0$ plants $\times$ population. There is an advantage in the evaluation of $S_1$ progeny itself since it provides the basis for a more accurate evaluation than the phenotypic appearance of individual $S_0$ plants. Results obtained by Lonnquist (1949) indicated a significant change in yield after one cycle of selection. Synthetic-2 of the first cycle of selection for high and low yields produced 142 and 85%, respectively, of the parental variety in the first year test and 118 and 88% in the second year. In the following year, after visual mass selection, the low-yield Synthetic-3 equaled the parental variety, while the high-yield Synthetic-3 was 27% greater in yield. Lonnquist (1951) reported that the high- and low-yielding synthetics obtained by recurrent selection, using the parental population as tester, were also differentiated in relation to combining ability with the single-cross WF9 $\times$ M14. Lonnquist and McGill (1956) selected $S_1$ lines based on testcrosses with the parental variety in three populations (Krug, Reid, and Dawes open-pollinated varieties) and obtained increases in yield of 22, 9, and 9%, respectively. In two cycles of selection the relative performance of the four improved populations in percent of the double-cross hybrid US13 were 87, 86, 85, and 72% in the first cycle and 98, 95, 102, and 88% in the second cycle for the populations Krug, Reid, and Synthetics A and B, respectively. Crosses within each cycle in all possible combinations averaged 68.0 q/ha with 18.8% moisture in the first cycle and 71.9 q/ha with 19.2% moisture in the second cycle. The commercial hybrid US13 yielded 68.0 q/ha with 19.0% moisture. Lonnquist and Gardner (1961) showed that one cycle of selection resulted in an increase from 57.3 to 60.4 q/ha and from 54.3 to 56.7 q/ha in the populations Krug and Nubold Reid, respectively. There was also an increase in the cross between the two populations from 59.1 to 67.1 q/ha. Heterosis relative to mid-parent also increased from 6.0 to 14.6%, showing that the improvement of the populations was based mainly on additive effects.

Lonnquist (1952) conducted two cycles of recurrent selection in two populations using non-related synthetic varieties as testers. In the first and second cycles the Krug variety yielded 82.4 and 88.4% of the double-cross US13, respectively, followed by a slight increase in stalk lodging and grain moisture. Similarly, Synthetic A yielded 81.4 and 99.5% of US13 in the first and second cycles, respectively, followed by a slight increase in grain moisture but no apparent change in stalk lodging. Lonnquist (1961) reported that three cycles of selection in five populations (Krug,
Reid, Synthetics A and B, and SSS) resulted in an average yield increase of 7.2 q/ha, whereas the intercross within cycles also increased yield 7.2 q/ha. In another comparison involving two cycles in three populations (Krug and Synthetics A and B), an average increase of 3.9 q/ha was obtained. After four cycles an average increase of 3.4 q/ha per cycle was observed for the three populations; the intercrosses showed an average increase of 3.5 q/ha (Lonnquist, 1963). Thompson and Harvey (1960) reported that the mean yield of the testcrosses increased from 77.8 to 98.8% relative to check means after five cycles of recurrent selection in Synthetic A. Eberhart et al. (1967) used the procedure described by Lonnquist (1949) with the variety Kitale. An evaluation of this experiment was presented by Eberhart and Harrison (1973). Two cycles of selection resulted in 5.2 q/ha (15%) increase in yield at a level of 34.6 q/ha. There were, however, large differences between yield levels expressed across environments. In better environments the original Kitale yielded 65 q/ha, and the improved material yielded 74.4 q/ha (a 14% improvement) while in poor environments, where the original variety yielded only 15 q/ha, the improved material was expected to yield 2.4 q/ha (16%) higher.

Hull (1945) proposed the use of inbred lines and hybrids as testers. Therefore, recurrent selection for specific combining ability is selection based on testcrosses with a narrow base tester. Sprague and Russell (1957) selected two populations (Lancaster and Kolkmeier) to study selection for combining ability with a specific tester (inbred line Hy). The first cycle showed a gain of about 6% in the population × tester performance for each population and in the cross between the two populations. After two cycles of half-sib selection with a narrow base tester the gains per cycle, via the linear regression coefficient, were 4.3, 12.8, and 15.1% for the crosses, Lancaster × Hy, Kolkmeier × Hy, and Lancaster × Kolkmeier, respectively. Sprague et al. (1959) reported that selection for specific combining ability with the inbred line Hy gave an increase in testcross yields of about 4.2 and 14.7% per cycle, respectively, relative to the original population testcross. Yields of the populations themselves decreased slightly, but the cross between the two populations increased at a rate of 7.9% per cycle. Lonnquist (1961) showed the results from selection in Krug using the single-cross WF9 × M14 as tester. After three cycles the data indicated a gain per cycle (regression coefficient) of about 3.6% with a slight decrease in grain moisture and a slight increase in stalk lodging.

Penny (1959) and Penny et al. (1962) obtained significant progress in selecting two populations for combining ability with inbred tester B14. After two cycles of selection for high and low yields, changes in yield of the populations themselves were 7.1 and 4.3% per cycle after selection for high yield and −17.4 and −7.1% after selection for low yield for Alph and (WF9 × B7), respectively. The cross Alph × B14 changed about 7.2% per cycle for high yield, but selection for low yield also increased slightly (1% per cycle). The cross (WF9 × B7) × B14 showed a decrease when selected for low yield (−3.5% per cycle) but a very small increase for high yield (1.1% per cycle). The cross between the two populations indicated an increase (3.8% per cycle) in high × high and a decrease of 7.1% per cycle in low × low. When all data were considered the predominant type of selection appeared to have been for genes exhibiting partial to complete dominance and largely additive
7.2 Improvement from Intra-population Selection

Additional information was generated in Alph and (WF9 × B7) for combining ability with B14 (Russell et al., 1973). After five cycles of selection, progress from selection was by changes in the yield of the testcrosses [Alph × B14 and (WF9 × B7) × B14], the populations themselves, and in crosses with related and unrelated testers. The gains per cycle are shown in Table 7.11.

Table 7.11  Linear regression coefficients (gain per cycle for five cycles) in percent of the original mean of populations themselves and in crosses with related and unrelated testers

<table>
<thead>
<tr>
<th>Population</th>
<th>Yield (q/ha)a</th>
<th>Ears per100 plants</th>
<th>Moisture (%)b</th>
<th>Plant height (cm)b</th>
<th>Ear height (cm)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alph per se</td>
<td>5.4</td>
<td>3.4</td>
<td>2.4</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>Alph CO</td>
<td>6.0</td>
<td>2.9</td>
<td>1.4</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>B14</td>
<td>4.4</td>
<td>2.0</td>
<td>0.9</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>BSBB</td>
<td>6.7</td>
<td>3.2</td>
<td>1.3</td>
<td>1.7</td>
<td>2.3</td>
</tr>
<tr>
<td>(WF9 × B7) per se</td>
<td>3.8</td>
<td>4.1</td>
<td>1.3</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Alph CO</td>
<td>3.6</td>
<td>0.9</td>
<td>−0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>B14</td>
<td>1.8</td>
<td>−0.8</td>
<td>−1.0</td>
<td>0.6</td>
<td>1.3</td>
</tr>
<tr>
<td>BSBB</td>
<td>2.4</td>
<td>0.8</td>
<td>−0.2</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

aIn percent of the observed original mean
bIn percent of the estimated original mean
Source: Adapted from Russell et al. (1973)

Table 7.11 shows that several traits had an increasing trend in most instances. The population cross had an increase of 4.1 q/ha per cycle. Because the gain from selection was significant for the specific tester B14 as well as for unrelated testers, it was concluded that selection was mainly for general combining ability (mostly additive) effects.

Less pronounced selection effects were observed by Walejko and Russell (1977) in evaluating five cycles of selection in the populations Kolkmeier and Lancaster for combining ability with the inbred line Hy (Table 7.12).

Yield was the primary trait in selection, and many of the C0 vs. C5 comparisons for yield showed significant differences. Despite the small increase or decrease in observed yield in the populations themselves, it was concluded that the recurrent selection program was successful in increasing frequencies of genes affecting yield because the testcrosses with the inbred line Hy and the population cross showed significant changes in yield. They concluded that recurrent selection from use of an inbred line as tester seems to be an efficient method for improving breeding populations and that the tester can be replaced by another line without any loss of improvement from the previous tester.

Horner et al. (1976) reached to the same conclusion in a population selecting for combining ability with the single cross (F44 × F6) as tester. After seven cycles of selection the testcross performance showed 18% more grain yield, 9% lower ear height, and 35% less lodging. Performance was similar when the same selected populations were crossed with an unrelated synthetic, showing that the narrow base tester was effective for improving general as well as specific combining ability.
Table 7.12 Estimated mean (a) and gain per cycle (b: linear regression coefficient) for two populations, their cross, and in crosses with the specific tester Hy (Walejko and Russell, 1977)

<table>
<thead>
<tr>
<th>Population</th>
<th>Yield (q/ha)</th>
<th>Moisture (%)</th>
<th>Root</th>
<th>Stalk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Kolkmeier (K)</td>
<td>44.0</td>
<td>1.6</td>
<td>28.6</td>
<td>−0.3</td>
</tr>
<tr>
<td>Lancaster (L)</td>
<td>47.7</td>
<td>−2.0</td>
<td>22.6</td>
<td>−0.5</td>
</tr>
<tr>
<td>K × L</td>
<td>59.9</td>
<td>4.1</td>
<td>25.2</td>
<td>−0.8</td>
</tr>
<tr>
<td>K × Hy</td>
<td>63.2</td>
<td>4.4</td>
<td>26.7</td>
<td>−1.0</td>
</tr>
<tr>
<td>L × Hy</td>
<td>67.2</td>
<td>3.3</td>
<td>24.5</td>
<td>−0.9</td>
</tr>
</tbody>
</table>

*In percent of the estimated mean (a)

Eberhart et al. (1973) reported on seven cycles of selection in Iowa Stiff Stalk Synthetic using the double-cross IA13 as tester. Increase in yield in the testcross was linear at a rate of 1.65 q/ha (2.6% per cycle) but the population itself increased at a lower rate (1.4% per cycle).

Few negative results have been reported for half-sib selection with the use of testers. Selection for earworm resistance was shown to be ineffective by Widstrom et al. (1970), as indicated by the linear increase in earworm rating ($b = 0.07 \pm 0.05$) after four cycles of selection in a synthetic using a single-cross hybrid as tester.

Additional types of half-sib selection were reported. Moll (1959) proposed a procedure in which several plants in a population were used as males and each was crossed with four plants as females; the best half-sib families that identified the best male plants were recombined. One, two, and three cycles of selection in the populations Indian Chief, Jarvis, and (C121 × NC7)F₂ resulted in 11, 7, and 17% yield increase, respectively, over the original populations.

A summary of results from recurrent selection for general and specific combining ability for yield is shown in Table 7.13.

### 7.2.3 Full-Sib Family Selection (FS)

It seems full-sib family selection for intra-population improvement should have received greater attention than other methods because selection among full-sib families includes, in theory, selection for both additive and dominance genetic effects. Full-sib families are obtained from plant-to-plant crosses, thus providing better parentage control for selection of superior families but more time is needed for pollination and bag identification. If populations have one ear per plant, full-sib progenies will be made of a bulk of two ears. Therefore, amount of testing is limited by ear size. Plant-to-plant crosses were used by A. E. Blount as early as in 1868, as reported in the 1936 USDA Yearbook. Blount reported on 100 to 300 desirable plants that were pollinated with pollen from other desirable plants selected as males. Apparently there was no further test of the paired crosses, so the mixture of
<table>
<thead>
<tr>
<th>Population</th>
<th>Tester&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. of cycles</th>
<th>Population per se&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Testcross&lt;sup&gt;b&lt;/sup&gt;</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krug (high)</td>
<td>B</td>
<td>1</td>
<td>17.7</td>
<td>—</td>
<td>Lonnquist (1949)</td>
</tr>
<tr>
<td>Krug (low)</td>
<td>B</td>
<td>1</td>
<td>17.7</td>
<td>—</td>
<td>Lonnquist (1952)</td>
</tr>
<tr>
<td>Krug</td>
<td>B</td>
<td>2</td>
<td>6.8 (1)</td>
<td>—</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Krug</td>
<td>B</td>
<td>1</td>
<td>22.0</td>
<td>—</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Reid</td>
<td>B</td>
<td>1</td>
<td>9.0</td>
<td>—</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Dawes</td>
<td>B</td>
<td>1</td>
<td>9.0</td>
<td>—</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Krug</td>
<td>B</td>
<td>2</td>
<td>11.0 (1)</td>
<td>4.8</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Reid</td>
<td>B</td>
<td>2</td>
<td>9.0 (1)</td>
<td>1.5</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Syn. A</td>
<td>B</td>
<td>2</td>
<td>17.0 (1)</td>
<td>5.5</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Syn. B</td>
<td>B</td>
<td>2</td>
<td>16.0 (1)</td>
<td>11.3</td>
<td>Lonnquist and McGill (1956)</td>
</tr>
<tr>
<td>Lancaster</td>
<td>N</td>
<td>2</td>
<td>—</td>
<td>4.3</td>
<td>Sprague and Russell (1957)</td>
</tr>
<tr>
<td>Kolkmeier</td>
<td>N</td>
<td>2</td>
<td>—</td>
<td>12.8</td>
<td>Sprague and Russell (1957)</td>
</tr>
<tr>
<td>Lancaster</td>
<td>N</td>
<td>2</td>
<td>—</td>
<td>4.3</td>
<td>Sprague et al. (1959)</td>
</tr>
<tr>
<td>Kolkmeier</td>
<td>N</td>
<td>2</td>
<td>—</td>
<td>14.7</td>
<td>Sprague et al. (1959)</td>
</tr>
<tr>
<td>Alph</td>
<td>N</td>
<td>2</td>
<td>2.8</td>
<td>6.3</td>
<td>Penny (1959)</td>
</tr>
<tr>
<td>(WF9 × B7)F2</td>
<td>N</td>
<td>2</td>
<td>10.3</td>
<td>1.3</td>
<td>Penny (1959)</td>
</tr>
<tr>
<td>SS Syn.</td>
<td>N</td>
<td>4</td>
<td>—</td>
<td>1.3</td>
<td>Penny (1959)</td>
</tr>
<tr>
<td>Krug</td>
<td>B</td>
<td>3</td>
<td>3.9 (3)</td>
<td>4.9</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Reid</td>
<td>B</td>
<td>3</td>
<td>8.6 (3)</td>
<td>4.9</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Syn. A</td>
<td>B</td>
<td>3</td>
<td>0.5 (3)</td>
<td>4.1</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Syn. B</td>
<td>B</td>
<td>3</td>
<td>9.0 (3)</td>
<td>8.0</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>SS Syn.</td>
<td>B</td>
<td>3</td>
<td>13.9 (3)</td>
<td>6.2</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Krug</td>
<td>N</td>
<td>3</td>
<td>—</td>
<td>3.6</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Krug</td>
<td>B</td>
<td>1</td>
<td>5.4</td>
<td>8.9</td>
<td>Lonnquist and Gardner (1961)</td>
</tr>
<tr>
<td>Nubold</td>
<td>B</td>
<td>1</td>
<td>4.5</td>
<td>4.0</td>
<td>Lonnquist and Gardner (1961)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Population tester.

<sup>b</sup> Gain per cycle (in parenthesis).
<table>
<thead>
<tr>
<th>Population</th>
<th>Tester&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. of cycles</th>
<th>Population per se&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Testcross&lt;sup&gt;b&lt;/sup&gt;</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alph (high)</td>
<td>N</td>
<td>2</td>
<td>7.2</td>
<td>7.2</td>
<td>Penny et al. (1962)</td>
</tr>
<tr>
<td>Alph (low)</td>
<td>N</td>
<td>2</td>
<td>-17.7</td>
<td>1.0</td>
<td>Penny et al. (1962)</td>
</tr>
<tr>
<td>(WF9 × B7) (high)</td>
<td>N</td>
<td>2</td>
<td>4.4</td>
<td>1.1</td>
<td>Penny et al. (1962)</td>
</tr>
<tr>
<td>(WF9 × B7) (low)</td>
<td>N</td>
<td>2</td>
<td>-7.3</td>
<td>-3.5</td>
<td>Penny et al. (1962)</td>
</tr>
<tr>
<td>3 populations B</td>
<td>B</td>
<td>4</td>
<td>6.4 (5)</td>
<td>5.2 (2)</td>
<td>Lonnquist (1963)</td>
</tr>
<tr>
<td>Kitale</td>
<td>B</td>
<td>2</td>
<td>7.5</td>
<td></td>
<td>Eberhart and Harrison (1973)</td>
</tr>
<tr>
<td>Alph</td>
<td>N</td>
<td>5</td>
<td>5.9</td>
<td>4.6</td>
<td>Russell et al. (1973)</td>
</tr>
<tr>
<td>WF9 × B7</td>
<td>N</td>
<td>5</td>
<td>4.4</td>
<td>1.8</td>
<td>Russell et al. (1973)</td>
</tr>
<tr>
<td>Lancaster</td>
<td>N</td>
<td>5</td>
<td>-2.0</td>
<td>3.3</td>
<td>Walejko and Russell (1977)</td>
</tr>
<tr>
<td>Kolkmeier</td>
<td>N</td>
<td>5</td>
<td>1.6</td>
<td>4.4</td>
<td>Walejko and Russell (1977)</td>
</tr>
<tr>
<td>FSB (HT)</td>
<td>N</td>
<td>5</td>
<td>-</td>
<td>3.5 (6)</td>
<td>Horner et al. (1976)</td>
</tr>
</tbody>
</table>

<sup>a</sup>B: broad base, N: narrow base

<sup>b</sup>(1): From first to second cycle in percent of US13; (2): average performance in all possible cross combinations within cycles; (3): from first to third cycle in percent of estimated original mean; (4): average performance in a 12 × 12 diallel cross; (5): average gain (linear regression) in three populations (Krug and Synthetics A and B) in percent of estimated original mean; (6): C7 yields 3.3% in excess over C5
all crosses was similar to phenotypic (mass) selection with control of the pollen parent. The Blount White Prolific variety resulted from this selection and was widely distributed.

Harland (1946) suggested a method by which paired crosses were tested in replicated trials followed by selection of the best ones and recombination in detasseling blocks. Using this procedure with a local variety never selected before, an improved variety yielding 62.8 q/ha was reported. Lonnquist (1961) reported that selection of $S_0 \times S_0$ crosses in Krug III Synthetic resulted in an increase of 3.5% for high yield and a decrease of 6.1% for low yield (relative to the parental population) in 1 year of testing. In the following year/s test the performances of the first cycle for high and low yields were 114 and 101% of the parental population, respectively. The average of 2 years of testing indicated an increase of 8.9% for high-yield selection and a decrease of 2.6% for low-yield selection. At the same time a system of paired crosses was proposed; each $S_0$ plant was tested with two other genotypes instead of one. The system was called *chain crossing* and the procedure is in sequence as follows:

$$1 \times 2, 2 \times 3, \ldots, (n-1) \times n, n \times 1.$$  

It also was suggested to select plants on the basis of their average effect in the two crosses and then to select the best cross within each pair selected in the first phase. Selection for high and low yield in Krug III gave 10.6% increase and 4.9% decrease, respectively, after one cycle of selection evaluated in 2 years. Selection in the chain series, therefore, seemed to be more effective than the paired cross to improve yield.

Robinson and Comstock (1955) showed that one cycle of selection in four populations (CI21 × NC27, NC34 × NC45, Jarvis, and Weekly) resulted in an average yield increase of 9.8% over check yields. An additional cycle in the first two populations did not show any further increase. Moll and Robinson (1966, 1967) reported that three cycles of full-sib selection resulted in yield increases of 3.6 and 2.1% in the populations Jarvis and Indian Chief across 2 years of testing. The cross between the two populations increased 3.8% in yield, as indicated by 1 year of testing. Moll and Stuber (1971) obtained significant increases in yield after six cycles of full-sib family selection. The gain per cycle, in grams per plant, was 9.47, 6.81, and 9.76 for the populations Jarvis, Indian Chief, and Jarvis × Indian Chief, respectively. At a planting density of 50,000 plants/ha, those increases correspond to 4.74, 3.40, and 4.88 q/ha. The variety cross yield increased at a rate of 7.21 g per plant (3.60 q/ha). Moll et al. (1975) evaluated five criteria for one cycle of divergent selection through full-sib families from remnant seed for recombination in the variety Jarvis. The average effects of selection for each criterion are shown in Table 7.14. Data show that selection resulted in significant differences from the original variety for both yield and ear height for all selection criteria employed.

Four cycles of selection among full-sib families were completed by Jinahyon and Moore (1973). They observed that yield increased at a rate (regression coefficient) of about 7.9% per cycle while plant lodging decreased from 24 to 8% across
Table 7.14  Average yield, plant height, and ear height of subpopulations resulting from five selection criterion

<table>
<thead>
<tr>
<th>Selection criteriona</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Jarvis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jarvis</td>
</tr>
<tr>
<td>Yield (g/plant)</td>
<td>249.5</td>
<td>234.5</td>
<td>247.2</td>
<td>249.5</td>
<td>250.4</td>
<td>239.5</td>
</tr>
<tr>
<td>Plant height (cm)</td>
<td>284.5</td>
<td>264.1</td>
<td>271.8</td>
<td>285.7</td>
<td>293.6</td>
<td>282.5</td>
</tr>
<tr>
<td>Ear height (cm)</td>
<td>126.5</td>
<td>110.2</td>
<td>115.3</td>
<td>128.5</td>
<td>136.1</td>
<td>124.5</td>
</tr>
</tbody>
</table>

aA, B: single-trait selection for yield and ear height, respectively; C, D: restricted selection index with desired change of −5.1 and +5.1 cm in ear height, respectively; E: selection index for maximum expected change in yield

Source: Adapted from Moll et al. (1975)

cycles with a trend for reduce plant and ear heights. In an F<sub>2</sub> population, four cycles of full-sib family selection resulted in an increase in yield of 34% or a gain per cycle (regression coefficient) of 8.6% over the parental population (Genter, 1976a). Equally effective selection was reported by Compton (1977) in the variety Krug. After four cycles of full-sib family selection by use of a selection index called “harvestable yield” (index = yield × undropped ears × upright plants), the direct response of selection was successful. Yield increased from 64.1 to 72.9 q/ha (b = 2.9 ± 1.2 q/ha per cycle) and the index value increased from 49.2 to 59.7 q/ha (b = 2.8 ± 0.6 q/ha per cycle); percentage of undropped ears and upright plants increased and moisture decreased slightly.

Moll and Hanson (1984) conducted full-sib family selection within the open-pollinated varieties Jarvis and Indian Chief and the variety cross, Jarvis × Indian Chief. After 8–10 cycles of full-sib family selection, they reported grain yield increases of 26.2% for Jarvis, 5.6% for Indian Chief, and 20.6% for the Jarvis × Indian Chief cross. The gain realized in Jarvis was also reflected in the variety cross.

The CIMMYT maize breeding program has used full-sib family more extensively than other programs to develop improved varieties for use by producers in lesser developed areas (Vasal et al., 1982; Pandey and Gardner, 1992; Dowswell et al., 1996). One example is the selection for grain yield, days to silk, and plant height for eight tropical varieties (CIMMYT, 1984). After four to five cycles of full-sib selection, average responses per cycle of selection were increased yield (5.9%), fewer days to silk (−2.6 %), and reduced ear height (−4.6%). The CIMMYT maize breeders also have used within full-sib family as well as among full-sib family selection for different plant traits and pest resistance.

NDSAB(MER-FS)C13 was identified after 12 cycles of modified ear-to-row and 1 cycle of full-sib recurrent selection. Sixteen families out of 196 full-sib progenies were selected on the basis of a heritability index that included grain yield, grain moisture at harvest, root lodging, and stalk lodging. Selected families were recombined through a bulk-entry method that required only 32 rows in the winter nursery. The improved population was identified after testing populations per se
across 15 North Dakota (ND) environments and population crosses across 10 ND
environments. In crosses with BS21(R)C7 grain yield, root lodging, and stalk lodg-
ing were not statistically different from the top commercial transgenic check of a
major company.

The gains from full-sib family selection are similar to those of other recurrent
selection schemes. A summary of results from full-sib family selection is shown in
Table 7.15. Hallauer and Carena (2009) briefly stated that the North Dakota State
University (NDSU) maize breeding program has several ongoing full-sib recurrent
selection programs (see Table 7.25).

<table>
<thead>
<tr>
<th>Population</th>
<th>Number of cycles</th>
<th>Gain per cycle (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krug (high)a</td>
<td>1</td>
<td>8.9</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Krug (low)a</td>
<td>1</td>
<td>−2.6</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Krug (high)b</td>
<td>1</td>
<td>10.6</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Krug (low)b</td>
<td>1</td>
<td>−4.9</td>
<td>Lonnquist (1961)</td>
</tr>
<tr>
<td>Jarvis</td>
<td>6</td>
<td>3.5</td>
<td>Moll and Stuber (1971)</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>6</td>
<td>2.8</td>
<td>Moll and Stuber (1971)</td>
</tr>
<tr>
<td>(Jarvis × Indian Chief)c</td>
<td>6</td>
<td>2.5</td>
<td>Moll and Stuber (1971)</td>
</tr>
<tr>
<td>(Jarvis × Indian Chief)-syn.</td>
<td>6</td>
<td>2.8</td>
<td>Moll and Stuber (1971)</td>
</tr>
<tr>
<td>(CI21 × NC7)</td>
<td>10</td>
<td>4.0</td>
<td>Moll and Stuber (1971)</td>
</tr>
<tr>
<td>Cupurico × Flint Composite</td>
<td>4</td>
<td>7.1</td>
<td>Jinahyon and Moore (1973)</td>
</tr>
<tr>
<td>(Va17 × Va29)F2</td>
<td>4</td>
<td>9.3</td>
<td>Genter (1976a)</td>
</tr>
<tr>
<td>CIMMYT</td>
<td>4–5</td>
<td>5.9</td>
<td>CIMMYT (1984)</td>
</tr>
<tr>
<td>NDSAB(MER-FS)</td>
<td>1</td>
<td>12.5</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>NDBS22(R-FS)d</td>
<td>2</td>
<td>15.2</td>
<td>Sezegen and Carena (2009)</td>
</tr>
<tr>
<td>Krugc</td>
<td>4</td>
<td>4.7(5.8)</td>
<td>Compton (1977)</td>
</tr>
</tbody>
</table>

*aDivergent selection of full-sib families obtained by paired crosses; *bdivergent selection of full-sib families obtained by chain crossing method; *ccross between populations over cycles of intra-
population improvement; *dIndirect response of divergent selection for cold tolerance; *egain per cycle on yield and on “index” (in parentheses), respectively; *fsingle-trait (yield) selection

7.2.4 Inbred Family Selection (S)

Recurrent selection breeding procedures, especially based on inbred progenies, have
been effective for germplasm enhancement and the development of useful sources
of inbred lines. Recurrent selection based on inbred progenies at different levels of
inbreeding have shown progress in cold tolerance, pest resistance, and grain qual-
ity but has been challenging for improving grain yield in the long term. Selfing
was first used in recurrent selection systems to maintain tested genotypes. The level
of inbreeding used is arbitrary, but usually either $S_1 (F = 0.5)$ or $S_2 (F = 0.75)$
progenies are used to reduce the length of each cycle of selection. The combina-
tion of $S_1$ and $S_2$ progenies if often used to screen visually among $S_1$ progenies
and reduce the number of entries for yield trials. S₁–S₂ multi-stage recurrent selection was designed to increase the value of agronomic traits while improving grain yield. To be successful, complementary selection between both stages of selection is essential.

The production of progenies is straightforward as few rows are needed for self-pollinating hundreds of plants. The basic scheme consists in generating S₁ or S₂ families, evaluating inbred families in replicated trials, and recombining remnant seed from selected progenies. The evaluation and recombination phases can be combined in the same season (Dhillon and Khehra, 1989). Even if a winter nursery is available and comparisons are made on estimates of genetic gain per year, however, there is no clear advantage over other recurrent selection methods. Direct effects of selection are measured through selfed populations while indirect effects are represented by populations per se (Hallauer, 1992). It has been shown that variability among progenies increases with inbreeding, so use of inbred families in a recurrent selection system is mainly for characters of low heritability. In addition, as a consequence of the increased additive genetic variance among inbred progenies heritability would tend to increase. Moreover, lower means of inbred progenies than non-inbred ones do not imply less precision.

Advantages of inbred progeny selection include greater heritability of traits, easy detection of desirable genotypes and development of inbred lines, and the elimination of deleterious alleles (Hallauer et al., 1988). Iglesias and Hallauer (1989) stated, however, that the selection against unfavorable alleles must be careful since many favorable exotic alleles may be linked with undesirable factors. Moreover, the fixation of favorable as well as unfavorable alleles may be important with small population sizes. Mulamba et al. (1983) reported a considerable reduction in genetic variability (55.2%) in BSK(S)C8 with respect to other selection methods suggesting fixation of loci and consequently less potential for improvement.

Hull (1945) suggested inbred progeny recurrent selection (based on S₁ progenies) as an alternative to half-sib recurrent selection. After phenotypic recurrent (mass) selection was shown to be effective to increase resistance to *H. turcicum* (Jenkins et al., 1954), S₁ family selection was used by Bojanowski (1967) in an attempt to increase resistance to smut (*U. maydis*). Selection in a selfing series also was conducted to provide a basis for comparison. However, results were disappointing relative to effectiveness both of continuous inbreeding and selection and of recurrent selection by means of S₁ families. Further reports have demonstrated the effectiveness of inbred family selection for improvement of quantitative traits.

Penny et al. (1967) selected five populations for European corn borer (ECB, *Ostrinia nubilalis* Hübner) resistance during three cycles. It was observed that two cycles of selection were sufficient to shift frequencies of resistance genes to a higher level in all varieties and that three cycles produced essentially borer-resistant varieties. Klenke et al. (1986) reported data after five cycles of selection, which supported the conclusions by Penny et al. (1967) for levels of resistance to corn borer infestation, but one serious correlated effect was significant grain yield reductions. After the single-gene Bt maize system was introduced by industry in 1996
practically no reports were published on increasing the frequency of resistance alleles to ECB.

Jinahyon and Russell (1969a, 1969b) evaluated three cycles of recurrent selection to improve stalk rot (Diplodia zea [Schw.] Lev.) resistance in the open-pollinated variety Lancaster. Progress of improvement was determined in the cycle populations per se, the population testcrosses with WF9 × Hy and os420 × 187–2, and the diallel crosses among C0, C1, C2, and C3 populations. All methods of evaluation based on artificial inoculation with D. zea showed significant improvement for stalk rot resistance. Correlated changes with selection for stalk rot resistance were greater plant vigor, later maturity, better disease resistance, greater stalk strength, and greater yields in hybrid crosses. Devey and Russell (1983) also evaluated later cycles of selection for better stalk rot resistance. Levels of resistance and stalk strength improved with further cycles of selection, but grain yield decreased with selection for greater tolerance to D. zea infection.

Scott and Rosenkranz (1974) used three variations of inbred family selection to increase resistance to corn stunt: (1) selecting the best plant in the best 10 of 463 S1 progeny rows, (2) selecting the best 23 progenies derived from the original 463 S1 progenies, and (3) selecting the best 10 of 100 S1 progenies that were evaluated in a replicated test. Each method of selection was effective, but the most effective was that based on replicated tests (method 3); after one cycle of selection 21% of the plants in the improved population were infected with corn stunt compared with 52% diseased plants in the original population.

Maize production has expanded into very short-season environments that require germination and growth in cooler environments. Therefore, an important goal for maize breeders is to improve local germplasm sources of inbred lines that are able to grow under very cold and wet/dry conditions with increased risk for reduced seedling emergence and growth. Hoard and Crosbie (1985) defined cold tolerance as the ability of a genotype to emerge from the soil and to grow vigorously after emergence in cold, wet soils, and cold air temperatures. Recurrent selection was proposed as one of the best approaches to develop superior cold-tolerant germplasm sources since it takes advantage of all possible gene contributions (Mock and Eberhart, 1972). Mock and Bakri (1976) obtained progress in selecting for cold tolerance based on S1 progenies. Percent of emergence and dry weight increased 30.1 and 13.2%, respectively, in the population BS13(SCT). On the other hand, those traits did not respond consistently to selection of S1 families in the population BSSS2(SCT). Emergence index was not changed by selection in either population. Hoard and Crosbie (1985) evaluated the same populations after five cycles of the same method and concluded that selection would be effective with the use of a selection index including emergence percentage across more than one environment. Sezegen and Carena (2009) conducted divergent selection for cold tolerance across known locations targeted at cold tolerance screening (e.g., northern North Dakota and Montana areas of higher elevation). Selection was based on a rank-summation index that included seedling vigor and emergence percentage at a rate of one cycle per year. S1 progenies were produced in winter nursery and trials evaluated across
locations in the following summer. Since selection was conducted before flowering recombination was performed during the same season in the summer nursery utilizing the ‘intra-diallel’ method (Sezegen and Carena, 2009). Direct response to selection was not significant for these traits across populations after evaluating only two cycles of divergent selection. However, many desirable correlated changes were obtained and long-term selection was encouraged before evaluation. In addition, inbred progeny recurrent selection for cold tolerance utilizing over 200 progenies under extreme cold environments (e.g., high elevations under cool controlled conditions) have been promising for inbred line development at NDSU.

Inbred progeny selection is the intra-population recurrent selection method that has, in theory, the advantages for the highest expected genetic gain for grain yield (Comstock, 1964), and it should be an efficient method of improving populations as a source of inbred lines (Smith, 1979b). Also, it seems that additive effects of alleles with partial to complete dominance account for greater portion of the genetic variability in genetically broad-based populations (Tanner and Smith, 1987, Hallauer, 1992). Based on this assumption, inbred progeny selection seems to be appropriate.

Inbred progeny selection has been successful for increasing yield in the short term. Jinahyon and Moore (1973) observed an increase in yield of 8.3% per cycle after two cycles of S1 family selection in Thai Composite; stalk lodging in C0 decreased from 53 to 17% in the second cycle. They also observed slight decreases in plant and ear height and no changes in the number of days to silk. Weyhrich et al. (1998) showed the superiority of S2 recurrent selection over five selection methods after four cycles of selection in BS11. Hallauer (1978) also reported a 3.1% per cycle increase after five selection cycles for BSK. Response to selection was also effective in Leaming and Midland Yellow Dent after three cycles of S1–S2 recurrent selection (Carena and Hallauer, 2001a). For grain yield, average increases per cycle of selection of 0.28 Mg/ha (10.1%) in Leaming and 0.22 Mg/ha (14.9%) in Midland Yellow Dent were highly significant ($P \leq 0.01$). Dominance genetic effects were responsible for the increase in the frequencies of favorable alleles in Leaming. Favorable correlated responses were observed in traits that are essential for adaptation. Root lodging decreased 7.0% per cycle in Midland Yellow Dent while stalk lodging decreased 3.8% per cycle ($P \leq 0.01$) in Leaming. In this case, S1–S2 recurrent selection was an effective method to adapt and improve both populations, suggesting an adequate balance between the selection among S1 and S2 progenies.

Inbred progeny selection was not as efficient as expected, especially for grain yield in the long term (Tanner and Smith, 1987; Iglesias and Hallauer, 1991; Helms et al., 1989; Lamkey, 1992). Significant improvements for grain yield are realized from inbred progeny selection for the first two to four cycles of selection and level off in later cycles. It seems maize populations are more responsive in the long term if crosses (half-sib or full-sib families) are made and evaluated. Maize is nearly 100% cross-pollinated and does not seem amenable to inbred progeny selection for the long term. Three populations were studied using the two-stage (S1–S2) procedure (Iglesias and Hallauer, 1989), and because these populations included exotic germplasm, emphasis was given to maturity in the S1 progeny stage. A common result was the lack of selection response for yield after their third cycle of
7.2 Improvement from Intra-population Selection

Selection responses in other domestic backgrounds have shown the same trend. Tanner and Smith (1987) have shown no selection response in BSK after four cycles of inbred progeny selection. Inbred progeny selection was not effective in BS13(HI)C7 after seven cycles of half-sib recurrent selection. Lamkey (1992) showed no selection response after four and six selection cycles in BS13. The same population showed no response after four cycles of selection (Helms et al., 1989). The lack of continuous improvement via inbred progeny recurrent selection was associated with significant decreases in genetic variability (Mulamba et al., 1983; Hallauer, 1992) as a consequence of small effective population sizes (Weyhrich et al., 1998), especially in early cycles where frequencies of desired alleles were low, and fixation of loci associated with selection. Inbreeding depression due to genetic drift, however, was reported to be lower with inbred progeny selection (Helms et al., 1989; Garay et al., 1996) and less important than with inter-population selection methods (Oyervides-García and Hallauer, 1986; Rodriguez and Hallauer, 1988). The direction and amount of dominance can vary from gene to gene (Kearsey, 1993) and can decrease the importance of additive effects. The importance of dominance variance in inbred progeny selection is usually minimized but the dominance variance ranges from 19 to 25%, values that should not be underestimated. An increased frequency of favorable alleles is the consequence of selection response. This can be achieved not only by changing additive effects but also with the complement of dominant effects, although effects are not always complementary (Stojšin and Kannenberg, 1994). Negative correlations between traits selected in different stages (e.g., S1–S2 selection) could be another cause of lack of response (Oyervides-García and Hallauer, 1986; Iglesias and Hallauer, 1989). Hallauer (1992) suggested that the reason of lacking response after two to three cycles of S1–S2 selection could be the elimination of high-yielding genotypes due to intense selection for traits, such as pest resistance, which may cause a severe competition for photosynthates.

Genetic improvement using S1 or S2 families has been used for a broad range of traits (Hallauer, 1992; Oyervides-García and Hallauer, 1986; Iglesias and Hallauer, 1989; Carena and Hallauer, 2001a). The combination of S1 or S2 selection on a progeny basis and evaluation of S1 or S2 testcross is a common feature of applied maize breeding programs. Inbred progenies that have superior GCA and SCA are retained in the breeding nursery for further inbreeding and selection and additional evaluation in crosses. Elite inbred lines can be used in pedigree breeding to develop recycled lines. This protocol is continued to enhance the combination of alleles important in hybrids. Hence, testing identifies superior inbred lines that can be recurrently selected to develop better inbred lines. Two-stage selection (S1 + S2 recurrent selection) has been a cost-effective methodology (Hallauer, 1992). In the first stage, traits with relatively higher heritability estimates than yield are evaluated among-and within-S1 progenies in the target environment based on large sample sizes while in the second stage selected S2 progenies are evaluated in replicated yield trials. Following testing remnant S1 seed of the superior S2 progenies is recombined.

Table 7.16 lists breeding programs that conducted inbred progeny selection for at least three selection cycles.
Table 7.16  Genetic gains in populations “per se” for grain yield from maize selection experiments based on inbred progeny selection

<table>
<thead>
<tr>
<th>Population</th>
<th>Cycles</th>
<th>Genetic gain ($\Delta G$)(^a)</th>
<th>%/cycle</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSK(S)C4</td>
<td>4</td>
<td>4.1</td>
<td>4.1</td>
<td>Burton et al. (1971)</td>
</tr>
<tr>
<td>BSK(S)C4</td>
<td>4</td>
<td>2.5</td>
<td>2.5</td>
<td>Genter and Eberhart (1974)</td>
</tr>
<tr>
<td>BSK(S)C8</td>
<td>8</td>
<td>2.9</td>
<td>2.9</td>
<td>Tanner and Smith (1987)</td>
</tr>
<tr>
<td>BSK(S)C8</td>
<td>8</td>
<td>4.0</td>
<td>4.0</td>
<td>Rodriguez and Hallauer (1988)</td>
</tr>
<tr>
<td>BS13(S)C4</td>
<td>4</td>
<td>1.3</td>
<td>1.3</td>
<td>Helms et al. (1989)</td>
</tr>
<tr>
<td>BS13(S)C6</td>
<td>6</td>
<td>0.2</td>
<td>0.2</td>
<td>Lamkey (1992)</td>
</tr>
<tr>
<td>BS13(S)C6</td>
<td>6</td>
<td>1.2</td>
<td>1.2</td>
<td>Holthaus and Lamkey (1995)</td>
</tr>
<tr>
<td>BS11(S)C5</td>
<td>5</td>
<td>2.2</td>
<td>2.2</td>
<td>Weyhrich et al. (1998)</td>
</tr>
<tr>
<td>BS11(S2)C4</td>
<td>4</td>
<td>4.8</td>
<td>4.8</td>
<td>Weyhrich et al. (1998)</td>
</tr>
<tr>
<td>BS16(S)C3</td>
<td>3</td>
<td>3.2</td>
<td>3.2</td>
<td>Rodriguez and Hallauer (1988)</td>
</tr>
<tr>
<td>BS16(S2)C4</td>
<td>5</td>
<td>−1.2</td>
<td>−1.2</td>
<td>Iglesias and Hallauer (1989)</td>
</tr>
<tr>
<td>BS2(S)C4</td>
<td>4</td>
<td>6.0</td>
<td>6.0</td>
<td>Rodriguez and Hallauer (1988)</td>
</tr>
<tr>
<td>BS2(S2)C5</td>
<td>5</td>
<td>4.5</td>
<td>4.5</td>
<td>Iglesias and Hallauer (1989)</td>
</tr>
<tr>
<td>VCBS(S)C4</td>
<td>4</td>
<td>5.4</td>
<td>5.4</td>
<td>Genter and Eberhart (1974)</td>
</tr>
<tr>
<td>BSSS2(S)C3</td>
<td>3</td>
<td>6.2</td>
<td>6.2</td>
<td>Oyervides-G and Hallauer (1986)</td>
</tr>
<tr>
<td>BSTL(S2)C5</td>
<td>5</td>
<td>5.2</td>
<td>5.2</td>
<td>Iglesias and Hallauer (1989)</td>
</tr>
<tr>
<td>Leaming(S1−S2)C3</td>
<td>3</td>
<td>10.1</td>
<td>10.1</td>
<td>Carena and Hallauer (2001a)</td>
</tr>
<tr>
<td>Midland(S1−S2)C3</td>
<td>3</td>
<td>14.1</td>
<td>14.1</td>
<td>Carena and Hallauer (2001a)</td>
</tr>
</tbody>
</table>

\(^a\)Values are based on $\Delta G = \left( \frac{C_n - C_0}{C_n} \right) \times 100$

7.2.5 Combination and Comparison of Intra-population Selection Methods

Inbred line performance is usually different from their derivative hybrids. Therefore, the modification of the inbred progeny selection methodology would seem reasonable. Mild selection among selfed progenies and yield evaluation of testcross seed from acceptable progenies has been proposed as a change needed to permit continued selection response (Goulas and Lonnquist, 1976; Moreno-González and Hallauer, 1982; Hallauer, 1992). The use of the inbred progeny selection based on the two-stage method followed by half-sib selection is being utilized. For instance, BS26 Lancaster population followed this system of genetic improvement at Iowa State University where 400 self-pollinated ears are planted ear-to-row for an initial screening for agronomic traits. Then, the approximately top 225 S\(_1\) progenies are crossed with a BSSS tester for multi-trait, multi-location trials of S\(_1\) × tester hybrids. In each cycle of selection, testers from the Iowa Stiff Stalk heterotic group have changed and response of selection has been successful.

The relative importance of selection for specific and for general combining abilities was studied by Lonnquist and Rumbaugh (1958). From the variety Krug
(KI), a sample of 152 S₀ plants was selected and testcrossed with the single-cross WF9 × M14. S₁ seed from the best 31 testcrosses was intermated to form the synthetic KII. Because KII showed no difference from KI, a new sample of 91 S₀ plants was selfed and testcrossed with the parental KI as tester. At the same time 30 S₁ lines from the first sample were testcrossed with KI to make 121 testcrosses available for yield testing. Sixteen of 121 testcrosses were selected and the KII(s) synthetic was obtained after recombination of remnant S₁ seeds. The second cycle synthetics, KII (based on narrow base tester) and KII(s) (based on broad base tester), were subsequently compared in performance trials. They yielded 95.0 and 98.5% of the double-cross US13, which yielded 62.6 q/ha. Thus selection for general combining ability apparently was more effective in selection of lines having greater additive genetic effects and produced a greater population yield than selection based on the narrower genetic base tester which would agree with the study of BS26 above. A different conclusion was presented by Horner et al. (1963) using a narrow base tester F6 and the parental broad base population Florida 767 as testers. Performance of testcrosses in percent of the double-cross check hybrid Dixie 18 was greater for the narrow base tester in the first three cycles but about 6% greater for the broad base tester in the fourth cycle; the linear rate of increase was 1.7 and 3.4% of Dixie 18 per cycle for the narrow base and broad base testers, respectively. Composites formed by intercrossing lines in the first three cycles were crossed with 11 different testers. In the F6 series grain production increased significantly from 96.3 to 102.8% relative to the grand mean, but no significant change (98.6–100.8%) was observed in the broad base tester series when crossed with the same testers. The combining ability of the two composites with the inbred line F6 showed an average gain in yield of 6.5% per cycle in the F6 series but only 1.5% in the broad base tester series. The conclusion was that recurrent selection for combining ability with an inbred line tester was more effective than with a broad base tester in improving grain yield in maize.

Horner et al. (1969) again compared efficiencies of inbred line F6 and a broad base tester (parental population) but also included selection based on S₂ families themselves for higher grain yields. After three cycles of selection, evaluation was based on (1) random-mated populations, (2) selfed populations (bulk of S₁ lines), and (3) crosses with 11 unrelated testers. Differences among methods were detected only in the second and third cycles. The highest yielding random-mated population was obtained using the parental population as tester, whereas the highest yielding selfed population was produced by selection of S₂ progenies themselves. The average combining ability with unrelated testers increased significantly (5.2%) for the three methods, but there were no significant differences among them. The inbred tester method was effective in increasing yield of selfed populations, but it was inferior to the other two methods for population improvement in overall evaluation. In addition, it was concluded that both the S₂ progeny method and the parental tester method were effective but in different ways. Selection based on inbred families places relatively more emphasis on contributions of homozygous loci while the parental tester method emphasizes the contribution of heterozygous loci. Using the methods described above, Horner et al. (1973) subsequently compared results
of five cycles of selection. General combining ability, evaluated through testcrosses with two broad base testers, showed a significant linear increase for all methods. Selection based on the inbred tester was significantly more effective than the two other methods, showing a gain in grain yield of 4.4% per cycle. The parental tester and S₂ progeny methods showed gains of 2.4 and 2.0%, respectively. Performance of the random-mated populations after adjustment for inbreeding depression also showed a linear increase, but differences among the three methods of selection were not significant. Such results and the result reported by Lonnquist (1968) suggest that selection based on inbred families (S₁ or S₂ lines) has been less effective than expected. The S₂ progeny method, particularly, is theoretically a more effective method for changing frequencies of genes having additive effects than are the testcross methods (Horner et al. 1969). After additional cycles of selection, Horner et al. (1989) compared selection response from inbred progeny and half-sib family selection. Response to half-sib family selection was greater than inbred progeny selection, which they interpreted because of response to selection for overdominant gene effects expressed in the testcrosses with the inbred tester.

Penny (1968) also compared narrow base (double cross) vs. broad base (synthetic) testers in a selection experiment with BSSS. After three cycles of selection involving the broad base tester and six cycles involving the narrow base tester, no difference was observed in gains per cycle, which were 1.4 and 1.8% per cycle, respectively. A composite formed from elite inbred lines showed an increase in yield of 5% over the original variety. However, time and effort required to develop inbred lines and recombining them causes this system not to be comparable with the former ones.

Lonnquist (1968) compared recurrent selection based on an unrelated tester (BIII synthetic) and based on the parental population (KIII synthetic) as tester. He also included selection based on S₁ lines themselves. An increase of 15% relative to KIII was observed in the derived population when the parental population (KIII) was used as tester. The population derived from selection of S₁ lines themselves showed an increase of 4% in grain yield, whereas no apparent gain resulted from selection based on testcrosses with an unrelated tester. In a previous experiment (Lonnquist and Lindsey, 1964), however, a comparison was made between S₁ recurrent selection and half-sib recurrent selection based on testcrosses with an unrelated population (BIII synthetic) as the tester. Results indicated that both evaluation procedures were effective with a slight advantage for the use of an unrelated tester.

Comparisons between S₁ line and testcross performance as a basis for selection have been presented in several other reports. Koble and Rinke (1963) found significant correlations between S₁ lines and testcrosses with a related and an unrelated tester for yield and several other traits. Genter and Alexander (1966) compared S₁ lines themselves and testcross performance as a basis for selection in the population CBS. Inbred lines obtained from the population selected on the basis of S₁ progeny performance showed an increase of 31.4% in yield after two cycles. However, yield of S₁ lines from the population selected on the basis of testcrosses (two unrelated single-cross testers) increased at a lower rate (17.9% in two cycles).
 Duclos and Crane (1968) evaluated one cycle of selection through \( S_1 \) line performance in testcross with three synthetics; \( S_1 \) lines from the original population and from the derived populations based on \( S_1 \) lines and on testcross (double cross as tester) performance yielded 31.9, 40.7, and 36.9\% of the checks. Testcrosses with unrelated testers yielded 85.7, 85.4, and 88.7\% of the checks, respectively. Results of one cycle of selection indicated that selection based on \( S_1 \) line performance produced the best yielding \( S_1 \) lines and selection based on testcross performance with a double-cross tester produced the best testcross yields with unrelated testers. Populations obtained by intermating top lines in each method showed that both methods resulted in highly significant yield improvement. The second cycle of selection, however, was not effective.

 Carangal et al. (1971) evaluated two cycles of selection based on two types of families. In the first cycle there was a superiority of 42.6\% of testcrosses over \( S_1 \) lines for yield, which is expected due to inbreeding. In the second cycle, yields of \( S_1 \) lines from selection based on \( S_1 \) performance were not different from those derived from selection based on testcross performance. However, yield of testcrosses in the population selected for combining ability (based on testcrosses) was slightly greater (2\%) than of testcrosses from selection based on \( S_1 \) performance. Advanced populations obtained by \( S_1 \) progeny and testcross performance exceeded the parental variety by 4.6 and 2.7\%, respectively.

 There has been a considerable effort to compare the progress of selfed progeny selection with half-sib family selection without any conclusive results in favor of one method or the other (Hallauer et al., 1988; Hallauer, 1992; Moreno-González and Cubero, 1993). There has not been, however, enough research in combining both methods. Burton et al. (1971) evaluated four cycles of selection in Krug Hi I Synthetic 3 based on \( S_1 \) progeny performance and on testcrosses with an unrelated double cross as tester. A bulk of \( S_1 \) lines from the populations derived by the two methods (\( S_1 \) and testcross) had better yield increases for the selfing series (38.7\%) than for the testcross series (12.0\%). Evaluation of selection effects in testcrosses with four single crosses showed that in the selfing series the increase in testcross performance was 10.6\% after four cycles of selection, whereas in testcross selection only 5.7\% yield increase was observed. The two selection studies were continued for four additional cycles and Tanner and Smith (1987) compared responses to \( S_1 - S_2 \) progeny and half-sib family selection after eight cycles of selection. Their results, after four cycles of selection, were similar to those reported by Burton et al. (1971). Response to inbred progeny selection was greater at the C4 compared with the C4 of half-sib family selection. At the C8, they found greater response to selection via half-sib family selection. Greatest grain yield was attained by the C4 via inbred progeny selection with no further gains with four additional cycles of selection. Crosses were made between populations developed from the two methods of recurrent selection with 7.1\% mid-parent heterosis for the C4 \( \times \) C4 cross and 14.1\% for the C8 \( \times \) C8 cross, suggesting different alleles were selected by the two methods. Inbreeding depression estimates showed 7.9\% less inbreeding depression in the C8 population from inbred progeny selection vs. the C8 population from half-sib family selection, suggesting that the allele frequencies for the two C8 populations were
different. Horner et al. (1989) reported similar results comparing inbred progeny and half-sib family recurrent selection in the Fla.767 maize population. They interpreted the different responses to selection for overdominant gene effects which were expressed in the half-sib families when crossed with the inbred tester F6. This supported Hull’s (1945) original suggestion that such comparisons of the two recurrent selection methods would provide evidence of the relative importance of the types of genetic effects important in maize crosses.

Genter (1973) reported results of two cycles of selection for yield in two populations; no progress was observed in either testcross (related or unrelated tester) or S1 selection in the VLE population. However, selection based on S1 performance was effective in improving VCBS population (14.3% over two cycles), but selection based on testcrosses showed a non-significant increase of 2.7% over the parental variety.

A comparison between full-sib family selection and S1 family selection (Silva and Lonnquist, 1968) showed that both selected populations were significantly higher yielding than the original variety Krug. The observed gains were 11.2 and 15.0% for S1 selection and full-sib selection, respectively, at selection intensities of 4 and 33%.

Thompson (1972) used three different methods in a divergent selection program for lodging resistance. To increase lodging resistance three cycles of testcross evaluation (cycles 1, 4, and 5), one cycle of S1 family selection (cycle 2), and three cycles of full-sib family selection (cycles 3, 6, and 7) were used. Selection to increase lodging was based on one cycle (first) of testcross evaluation followed by five cycles of full-sib family selection. Two populations (Synthetics 8 and 9) were used and selection progress was evaluated by crosses between comparable cycles of the two synthetics, e.g., second cycle (Synthetic 8) × second cycle (Synthetic 9). Compared with the cross of the original unselected synthetics as 100%, selection to increase and decrease lodging showed 28 and 228% as many erect plants, respectively, in the last cycle. Results of the first cycles of selection were previously reported by Thompson (1963).

Goulas and Lonnquist (1976) used combination selection in the sense that selection was based simultaneously on performance of two types of families, half-sibs and S1 progenies. By selecting on the mean performance of half-sib families and S1 progenies obtained from the same plants a significant improvement was obtained. After two cycles yield increased 24% over the parental variety. There was also a decrease in grain moisture (6% in two cycles) and an increase of 7% in ear height.

Moreno-Gonzalez and Hallauer (1982) have suggested that data from a combination of progenies may contribute to greater genetic response. The major concern would be what weights should be assigned to the information from the different types of progenies. But an index based on estimates of heritability for the different progenies could be used (Smith et al., 1981). If we assume we desire to improve a population by some combination of S1 progeny (to reduce frequencies of undesirable recessive alleles) and testcross (to test for combining ability), we can evaluate S1 progenies in one replication at three to four locations with one location in an isolation field. In the isolation field, the S1 progenies are detasseled and crossed with
a common tester (e.g., an elite inbred line from the opposite heterotic group or the source population). The S\textsubscript{1} testcrosses also would be evaluated in replicated trials. Data would be available for the S\textsubscript{1} progenies and the S\textsubscript{1} progeny testcrosses: final selections could be based on an index, such as

\[
\bar{X}_{ci} = \hat{h}_1^2 \bar{X}_{ij} + \hat{h}_2^2 \bar{X}_{tc},
\]

where

\(\bar{X}_{ci}\) are the final selections, \(\hat{h}_i^2\) the heritability estimates, \(\bar{X}_{ij}\) = S\textsubscript{1} means, \(\bar{X}_{tc}\), the testcross means.

Heritability estimates are calculated on a progeny mean basis from the ANOVA of S\textsubscript{1} progenies and S\textsubscript{1} testcrosses. The S\textsubscript{1} progenies and S\textsubscript{1} testcrosses will be grown in separate experiments, that may be grown in the same or separate environments the same year or different years. The final selections will not be affected by the possible options of data collection because quality of data are weighted by the respective heritability estimates of the S\textsubscript{1} and testcross means. The index can be extended to include more than one trait. There are several possible options for use of data for different types of progenies in recurrent selection. They would be particularly amenable to the schemes that have been suggested for modified ear-to-row selection (Lonnquist 1964; Compton and Comstock, 1976), modified S\textsubscript{1} progeny selection (Dhillon and Khehra, 1989), or any schemes that include developing inbred programs to produce testcrosses. The evaluation of two types of progenies would contribute to reducing inbreeding depression (inbred progenies) and determine relative combining ability of the inbred progenies (testcrosses).

A summary of results of several selection programs involving comparisons among the testcross and selfing methods are presented in Table 7.17.

### 7.3 Improvement from Inter-population Selection

Inter-population recurrent selection methods are used in crop species where hybrids are important and we want to exploit heterotic combinations. This set of methods has as goal to increase the difference in allele frequencies (e.g., see definition of heterosis), which would increase the rate of direct response in the cross. These have not been as popular as intra-population selection methods. Inter-population selection includes two initial populations, and direct response to selection is measured in the population cross. Responses in the two parental populations themselves are the indirect responses to selection. This should not be confused with intra-population selection where direct response would be either in the population itself or testcrosses.

Several selection programs for inter-population improvement were initiated after Comstock et al. (1949) proposed reciprocal recurrent selection (RRS). Common sense would consider including elite genetic materials with genetic divergence as choice of populations to improve. If knowledge of populations and their potential
Table 7.17 Effect of selection (gain per cycle) on yield of populations themselves and in testcrosses in comparative studies of recurrent selection

<table>
<thead>
<tr>
<th>Population</th>
<th>Type of family&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. of cycles</th>
<th>Population per se&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Testcross&lt;sup&gt;b&lt;/sup&gt;</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida 767</td>
<td>BT</td>
<td>4</td>
<td>—</td>
<td>3.4</td>
<td>Horner et al. (1963)</td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>4</td>
<td>—</td>
<td>1.7</td>
<td>Horner et al. (1963)</td>
</tr>
<tr>
<td>Krug III</td>
<td>BT</td>
<td>1</td>
<td>13.8</td>
<td>—</td>
<td>Lonnquist and Lindsey (1964)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1</td>
<td>10.9</td>
<td>—</td>
<td>Lonnquist and Lindsey (1964)</td>
</tr>
<tr>
<td>CBS</td>
<td>BT</td>
<td>2</td>
<td>9.0</td>
<td>(2)</td>
<td>Genter and Alexander (1966)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>2</td>
<td>15.7</td>
<td>(2)</td>
<td>Genter and Alexander (1966)</td>
</tr>
<tr>
<td>BSSS</td>
<td>BT</td>
<td>3</td>
<td>1.4</td>
<td>—</td>
<td>Penny (1968)</td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>6</td>
<td>1.8</td>
<td>—</td>
<td>Penny (1968)</td>
</tr>
<tr>
<td>Krug III</td>
<td>BT(P)</td>
<td>1</td>
<td>15.0</td>
<td>—</td>
<td>Lonnquist (1968)</td>
</tr>
<tr>
<td></td>
<td>BT(U)</td>
<td>1</td>
<td>4.0</td>
<td>—</td>
<td>Lonnquist (1968)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1</td>
<td>1.0</td>
<td>—</td>
<td>Lonnquist (1968)</td>
</tr>
<tr>
<td>Purdue Ex-syn. 1</td>
<td>BT</td>
<td>1</td>
<td>25.4</td>
<td>(2)</td>
<td>Duclos and Crane (1968)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1</td>
<td>38.7</td>
<td>(2)</td>
<td>Duclos and Crane (1968)</td>
</tr>
<tr>
<td>Krug Yellow Dent</td>
<td>FS</td>
<td>1</td>
<td>15.0</td>
<td>—</td>
<td>Silva and Lonnquist (1968)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1</td>
<td>11.0</td>
<td>—</td>
<td>Silva and Lonnquist (1968)</td>
</tr>
<tr>
<td>Composite</td>
<td>BT(P)</td>
<td>1</td>
<td>8.1</td>
<td>—</td>
<td>Horner et al. (1969)</td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>1</td>
<td>−3.2</td>
<td>—</td>
<td>Horner et al. (1969)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
<td>−2.6</td>
<td>—</td>
<td>Horner et al. (1969)</td>
</tr>
<tr>
<td>Syn. A</td>
<td>BT</td>
<td>2</td>
<td>2.7</td>
<td>—</td>
<td>Carangal et al. (1971)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>2</td>
<td>4.6</td>
<td>—</td>
<td>Carangal et al. (1971)</td>
</tr>
<tr>
<td>BSK</td>
<td>BT</td>
<td>4</td>
<td>3.9</td>
<td>—</td>
<td>Burton et al. (1971)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>4</td>
<td>1.9</td>
<td>—</td>
<td>Burton et al. (1971)</td>
</tr>
<tr>
<td>Florida 767</td>
<td>BT</td>
<td>4</td>
<td>—</td>
<td>16.7</td>
<td>Horner et al. (1973)</td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>4</td>
<td>—</td>
<td>19.6</td>
<td>Horner et al. (1973)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;2&lt;/sub&gt;</td>
<td>4</td>
<td>—</td>
<td>47.7</td>
<td>Horner et al. (1973)</td>
</tr>
<tr>
<td>VLE</td>
<td>BT</td>
<td>2</td>
<td>13.3</td>
<td>—</td>
<td>Genter (1973)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>2</td>
<td>1.3</td>
<td>—</td>
<td>Genter (1973)</td>
</tr>
<tr>
<td>VCBS</td>
<td>BT</td>
<td>2</td>
<td>1.4</td>
<td>—</td>
<td>Genter (1973)</td>
</tr>
<tr>
<td></td>
<td>S&lt;sub&gt;1&lt;/sub&gt;</td>
<td>2</td>
<td>6.7</td>
<td>—</td>
<td>Genter (1973)</td>
</tr>
</tbody>
</table>

<sup>a</sup>BT: broad base tester (P: parental; U: unrelated tester); NB: narrow base tester; S<sub>1</sub> and S<sub>2</sub>: selfed families; FS: full-sib families

<sup>b</sup>(1) In percent of Dixie-18; (2) evaluated through S<sub>1</sub> line performance; (3) testcross performance; (4) gain in testcross performance with Min 707 from first to second cycle in percent of first cycle; (5) performance in testcrosses; corresponding gains in selfed populations are 3.9 and 9.6% per cycle, respectively; (6) S<sub>2</sub> line performance per se; (7) in crosses with two unrelated testers the increases were 2.4, 4.4, and 2.0% per cycle, respectively.
in crosses is limited extensive data generation from mating designs (e.g., diallel) among elite populations would aid in the choice of populations to improve via RRS. RRS has been effective in selection of complimentary alleles because mid-parent heterosis increased, across populations, from 7.3% for C0 × C0 to 37.4% for Cn × Cn crosses (see Table 5, Hallauer and Carena, 2009).

A RRS program was initiated between the populations Iowa Stiff Stalk Synthetic (BSSS) and Corn Borer Synthetic No. 1 (BSCB1) at Iowa State University in 1949. After two cycles of selection Penny (1959) reported a gain of 5.1% per cycle direct response in the cross between the two populations. After four cycles of selection Penny (1968) reported that an increase in yield in the crossed population was from 60.8 to 65.2 q/ha after 2 years of testing (a gain of 1.8% per cycle). A least-squares analysis of several experiments presented by Penny and Eberhart (1971) indicated very little progress (1.2 q/ha) in the population cross. The parental population BSSS showed a small increase (1.4 q/ha per cycle) and BSCB1 a small decrease (−0.6 q/ha per cycle) in yield after three cycles of selection. Similar changes were observed by Hallauer (1970) in evaluating the same populations and population crosses. A more comprehensive evaluation after five cycles of selection was presented by Eberhart et al. (1973). An increase in yield of 2.73 q/ha (4.5%) per cycle was detected in the population cross. The population BSSS selected for combining ability with double-cross IA13 as tester also showed an increase in yield (2.31 q/ha or 3.8% per cycle) when crossed with corresponding cycles of BSCB1 from the RRS program. Changes in yields of the populations themselves were very small, i.e., 0.9 and 0.4% per cycle for BSCB1 and BSSS. There was also an increase in ears per 100 plants in the populations and population cross. Stalk lodging was reduced in nearly all improved strains and in their crosses. A slight decrease in plant height was observed in the populations and a slight increase in the population cross. Increase in yield through RRS in the population cross without improvement in the parental populations or in testcross to an unrelated tester was consistent with expectations from changing gene frequencies at loci involving overdominant (or pseudo-overdominant) gene action. Keeratinijakal and Lamkey (1993) evaluated RRS in BSSS and BSCB1 after 11 cycles of selection. Grain yield, the primary trait, increased 7.0% per cycle of selection in the population crosses (e.g., direct response). Correlated responses to selection for grain yield included smaller tassel size, more upright leaf orientation, and greater root and stalk strength. There were no consistent changes for grain moisture. For the populations themselves, grain yield increased for BSSS (2% per cycle) but no change in yield of BSCB1.

Results of an RRS program involving the varieties Jarvis and Indian Chief at North Carolina State University were reported by Moll (1959). After one cycle of selection mean yield of the testcrosses changed from 92 to 98% of the mean of commercial double crosses. Prediction for the second cycle suggested that testcross yields would be 7% greater than double-cross yields. Results of three cycles were subsequently reported by Moll and Robinson (1966, 1967). Increases in yield were 0.7, 0.2, and 2.9% per cycle in Jarvis, Indian Chief, and Jarvis × Indian Chief, respectively, in the second cycle. In the third cycle, greater increases in yield were detected in the populations per se (4.3 and 1.7% per cycle in Jarvis and Indian Chief,
respectively). However, a smaller increase (0.8% per cycle) was detected in the population crosses. A full-sib intra-population family selection program in the same varieties showed an increase of 3.6, 2.1, and 3.8% per cycle in Jarvis, Indian Chief, and Jarvis × Indian Chief, suggesting that in three cycles intra-population selection was more effective than RRS for improvement of the variety cross. However, the selection differential was greater in the former, so the effectiveness of selection was nearly equal in both methods. Heterosis did not seem to have been changed by the intra-population improvement, whereas an increase in heterosis was observed in the first two cycles in RRS followed by a decrease in the third cycle. Moll and Stuber (1971) evaluated six cycles of the same RRS program. Total changes in yield were 14, 7, and 21% for Jarvis, Indian Chief, and Jarvis × Indian Chief, respectively, over the original populations and population cross. Responses of both varieties to full-sib family selection were 2.1 times greater than their responses to RRS, but the variety hybrid in RRS showed 1.3 times greater response than in full-sib family selection. Accumulated selection responses in the variety composite after six cycles were 20.3% above the mean of the unselected populations, which is approximately the heterosis in the variety cross. Therefore, six cycles of selection were necessary in the variety composite to attain the yield level equivalent to the original variety cross, i.e., the starting point of RRS. In other words, six cycles were necessary to recover the loss of one-half the heterosis that is expected to occur in the composite of a variety cross. Evaluation of other traits showed that increases in yield were generally accompanied by decreases in plant and ear heights and increases in tillers and ears per plant. An additional evaluation was reported by Moll et al. (1978) for six and eight cycles of selection. Full-sib family selection was more effective than RRS for intra-population improvement, but RRS was more effective for the improvement of the variety cross. Moll and Hanson (1984) summarized the results for 10 cycles of RRS in Jarvis and Indian Chief. In this particular study, average direct response in the population crosses for grain yield was 2.7% per cycle after 10 cycles of RRS. Mid-parent heterosis increased from 6.6% for the C0 × C0 cross to 28.9% for the C10 × C10 cross.

A RRS program was initiated in Texas by J. S. Rogers. Results of the first two cycles of selection were reported by Collier (1959) and Douglas et al. (1961). Increases in yield were at rates of 10.0 and 1.8% per cycle in the parental populations Ferguson’s Yellow Dent and Yellow Surecropper. Yield of the population cross increased at a rate of 5.8% per year. Thompson and Harvey (1960) showed that average yield of testcrosses in an RRS program involving Synthetic A and Synthetic G increased from 82.5 to 94.0% of check yields from the first to the third cycle. Torregroza et al. (1972) obtained 4.5 and 15.0% gain per cycle in the populations Harinoso Mosquera and Rocamex V7, respectively, after two cycles of RRS. The population cross yielded 32 and 34% in excess over the original cross in the first and second cycles, respectively.

Results of two cycles of RRS in Kenya were reported by Darrah et al. (1972). The parental populations KII and Ec573 increased in yield at a rate of 0.6 and 3.0
7.3 Improvement from Inter-population Selection

q/ha per year, respectively. The population cross showed an increase of 3.3 q/ha per year. After three cycles Darrah et al. (1978) obtained a gain of 2.1 q/ha per year in the population cross and 1.0 q/ha per year in the parental population Ec573 but a very small change (−0.02 q/ha per year) in KII population.

Gevers (1974) reported results of three cycles of RRS in the populations Teko Yellow and Natal Yellow Horsetooth, using two methods for sampling the male parents before crossing. When the pollen parents were taken at random in the population, grain yield showed an increase of 7.5, 7.4, and 5.8% per cycle of selection in Teko Yellow, Natal Yellow Horsetooth, and Teko Yellow × Natal Yellow Horsetooth, respectively. When male parents were selected for agronomic traits, however, changes in yield were 7.1, −0.5, and 3.3%, respectively. It was concluded that random sampling of parents before testcross evaluation led to greater progress in the populations themselves and in the population cross than when parents were selected for agronomic traits. Heterosis increased at similar rates in both methods.

RRS has been used for improvement of traits other than yield. Jugenheimer and Bhatnagar (1962) reported results of two cycles of RRS for plant lodging. In one population, percent of erect plants shifted from 71 to 83% and 29% when selected for low and high lodging, respectively. In the other population, selection for low and high lodging resulted in changes from 62 to 81% and 32% erect plants, respectively. Thomas and Grissom (1961) evaluated two cycles of RRS in popcorn. The gains per cycle were 10.7% increase for popping volume, 6.9% increase in yield, and 37.8% decrease in root lodging where evaluations were in percent of check hybrids; RRS was effective in improving several traits in the crossed population.

Reciprocal full-sib selection (FS-RRS or FR) was first evaluated by Hallauer (1973). After one cycle of selection involving populations Iowa Two-Ear Synthetic (BS10) and Pioneer Two-Ear Composite (BS11), a gain of 10.1% was realized in the cross C1 × C1 over C0 × C0. After three cycles of selection Hallauer (1975a) reported that full-sib progenies were at a higher yield level than those from the original populations. Full-sib progenies of the original populations averaged four standard deviations below the mean of check hybrids. After one cycle of selection, full-sib progenies were only 1 standard deviation below the mean yield of six check hybrids. No improvement was observed in the second cycle, during which mechanical harvesting was used. After three cycles of reciprocal full-sib selection, yield increased 18, 17, and 9.7% in BS10, BS11, and BS10 × BS11, respectively (Hallauer, 1977). Eyherabide and Hallauer (1991) reported response to FS-RRS in BS10 and BS11 after eight cycles of selection. Average grain yield response was 7.5% per cycle in the population crosses with mid-parent heterosis increasing from 2.5% for the C0 × C0 cross to 39.6% for the C8 × C8 cross. The 60% increase in grain yield after eight cycles of FS-RRS was accomplished with correlated responses of reduced tassel size, more upright leaf orientation, earlier maturity, and greater root and stalk strength. Grain yield was the primary trait in
selection, but data for root and stalk quality, and maturity also were included for
the full-sib families considered in selection, with the $S_1$ progenies of the 20 full-sib
selections used for intermating within BS10 and BS11. Indirect responses within the
BS10 and BS11 populations themselves followed the same trends as the population
crosses (Eyherabide and Hallauer, 1991).

Paterniani (1971) proposed a modification of RRS (HS-RRS2, see Chapter 6).
It resulted in a 13.6% yield increase in the Dent Composite population and a 7.8%
decrease in the Flint Composite population after the first cycle of selection. The
population cross increased 3.7% in the first cycle and 6.4% over the original cross
after the second cycle (Paterniani, 1974c). Paterniani and Vencovsky (1978) sum-
marized reporting a 3.5% increase per cycle in the population cross after two cycles
of the HS-RRS2 program.

Paterniani and Vencovsky (1977) obtained a 7.5% gain in the cross between
populations Cateto and Piramex after one cycle of RRS based on testcrosses
of half-sib families (HS-RRS-1, see Chapter 6). A least-squares analysis involv-
ing the population means indicated that observed progress was significant and
deviations from the model were not significant. In addition, progress in the pop-
ulation cross (7.5%) was partitioned into two portions: 4.3% due to changes
in population means and 3.2% due to changes in heterosis in the population
cross.

A summary of results from inter-population improvement programs is presented
in Table 7.18. Hallauer and Carena (2009) updated and summarized data for RRS
studies where average grain yield response in the population cross per cycle of selec-
tion was 4.8%, ranging from 2.7 to 7.5%. The number of selection cycles ranged
from 1 to 11.

Hybrids produced from inbred lines are the most popular choice of culti-
vars used by maize producers within the major maize producing areas of the
world. If decisions are made to conduct long-term selection programs to genet-
ically enhance germplasm resources for both current and future use, it seems
reciprocal recurrent selection should be the appropriate selection method. Maize
breeding programs that emphasize development of inbred lines to produce hybrids
must consider either present or future heterotic groups (Chapter 10). Based on
the present available data, it seems RRS methods enhance the expression of het-
erosis of heterotic groups. However, very few ongoing RRS maize programs
remain.

Intra- and inter-populations long-term genetic improvement programs are also
essential if the goal is to utilize alternative lower cost maize production systems.
Maize population hybrids can exploit heterosis in a fashion similar to single-cross
hybrids through a population-hybrid concept (Carena, 2005). The concept is an
alternative for the utilization of genetically broad-based germplasm based on the
hybridization of elite populations to exploit heterosis. Carena (2005) presented
evidence that crosses between geographically isolated populations improved by
both intra- and inter-population selection programs also can identify population
hybrids with significantly greater heterosis expression (Table 7.19). In many cases
population hybrids obtained from improved elite populations have shown similar
Table 7.18  Effect of reciprocal recurrent selection on yield of populations themselves and on population cross

<table>
<thead>
<tr>
<th>Parental populations</th>
<th>Methoda</th>
<th>No. of cycles</th>
<th>Gain per cycle (%)</th>
<th>b</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
<td>A x B</td>
</tr>
<tr>
<td>Yellow Surecropper</td>
<td>Ferguson Y. Dent</td>
<td>HS-RRS</td>
<td>2</td>
<td>10.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Jarvis</td>
<td>Indian Chief</td>
<td>HS-RRS</td>
<td>3</td>
<td>4.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Stiff Stalk Syn.</td>
<td>Corn Borer Syn.</td>
<td>HS-RRS</td>
<td>4</td>
<td>1.4</td>
<td>—0.7</td>
</tr>
<tr>
<td>Jarvis</td>
<td>Indian Chief</td>
<td>HS-RRS</td>
<td>10</td>
<td>3.1</td>
<td>—0.7</td>
</tr>
<tr>
<td>Harinoso Mosquera</td>
<td>Rocamex V7</td>
<td>HS-RRS</td>
<td>2</td>
<td>4.5</td>
<td>15.0</td>
</tr>
<tr>
<td>Stiff Stalk Syn.</td>
<td>Corn Borer Syn.</td>
<td>HS-RRS</td>
<td>11</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Kitale II</td>
<td>Ecuador 573</td>
<td>HS-RRS</td>
<td>3</td>
<td>—0.1</td>
<td>5.0</td>
</tr>
<tr>
<td>Teko Yellow</td>
<td>N.Y. Horsetooth</td>
<td>HS-RRS</td>
<td>3</td>
<td>7.5</td>
<td>7.4</td>
</tr>
<tr>
<td>BS10</td>
<td>BS11</td>
<td>FS-RRS</td>
<td>8</td>
<td>3.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Dent Composite</td>
<td>Flint Composite</td>
<td>HS-RRS1</td>
<td>1</td>
<td>13.6</td>
<td>—7.8</td>
</tr>
<tr>
<td>Piramex</td>
<td>Cateto</td>
<td>HS-RRS1</td>
<td>1</td>
<td>3.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Dent Composite</td>
<td>Flint Composite</td>
<td>HS-RRS2</td>
<td>2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Jarvis</td>
<td>Indian Chief</td>
<td>HS</td>
<td>6</td>
<td>2.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Jarvis</td>
<td>Indian Chief</td>
<td>HS</td>
<td>8</td>
<td>2.5</td>
<td>1.4</td>
</tr>
</tbody>
</table>

a(1) From first to third cycle in percent of checks; (2) popcorn population selected for popping volume, yield, and lodging resistance; (3) percent of observed means; (4) for random sampling of male parents – gains were 7.4, −0.5, and 3.3% per cycle, respectively, when parents were selected for agronomic traits.
bSee Chapter 6
Table 7.19 Grain yield and agronomic performance obtained as a result of long-term recurrent selection programs conducted in elite temperate germplasm

<table>
<thead>
<tr>
<th>Pedigree</th>
<th>GY</th>
<th>RL</th>
<th>SL</th>
<th>E</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS10(FR)C8 × BS11(FR)C8</td>
<td>7.5</td>
<td>3.6</td>
<td>12.5</td>
<td>8</td>
<td>Eyherabide and Hallauer (1991a)</td>
</tr>
<tr>
<td>BS10(FR)C13 × BS11(FR)C13</td>
<td>6.5</td>
<td>0.9</td>
<td>28.9</td>
<td>7</td>
<td>Hallauer (unpublished)</td>
</tr>
<tr>
<td>BSSS(R)C11 × BSCB1(R)C11</td>
<td>6.8</td>
<td>5.7</td>
<td>11.4</td>
<td>7</td>
<td>Keeratinijakal and Lamkey (1993)</td>
</tr>
<tr>
<td>BS21(R)C7 × BS22(R)C7b</td>
<td>7.4</td>
<td>0.0</td>
<td>3.4</td>
<td>20</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × CGSS(S1–S2)C5</td>
<td>8.0</td>
<td>0.1</td>
<td>9.0</td>
<td>20</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × CGL(S1–S2)C5</td>
<td>7.7</td>
<td>1.4</td>
<td>6.9</td>
<td>20</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS22(R)C7 × LEAMING(S)C4</td>
<td>7.1</td>
<td>7.9</td>
<td>8.1</td>
<td>20</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × NDSAB(MER-FS)C13</td>
<td>7.2</td>
<td>0.4</td>
<td>7.4</td>
<td>10</td>
<td>Carena (unpublished)</td>
</tr>
</tbody>
</table>

*Pedigree of crosses, grain yield (GY), root lodging (RL), stalk lodging (SL), number of environments (E) used for evaluation, and references

bThis cross had significantly larger grain moisture at harvest for North Dakota conditions

Source: Adapted from Carena and Wicks III (2006)

performance statistically to commercial hybrids. In order to identify them strong pre-breeding breeding programs are encouraged.

The public sector developed the inbred–hybrid concept. Shull decided to stop his research in the mid-1910s due to the practical limitations of producing seed on inbred lines. The private sector was instrumental to the practical success of hybrid corn. Heterosis might have been a tool used to focus research on the public inbred–hybrid concept as the only mechanism for maize genetic improvement. As a consequence, efforts to improve populations and their crosses were minimal during the period from 1920 to 1950. Renewed interest in improving breeding populations redirected some of the breeding efforts toward improving elite and genetically broad-based populations via recurrent selection. These efforts included extensive improvement of populations per se but limited testing of elite population hybrids. Elite population hybrids are those crosses between divergent populations that have undergone long-term selection and are competitive at least in agronomics with commercial hybrids (Carena and Wicks III, 2006). Population crosses were not widely accepted because of poor choice of germplasm. Adequate choice of germplasm after extensive testing and utilization of desirable breeding methodologies for genetic improvement have shown excellent high-parent heterosis values (Table 7.20).

Table 7.21 shows recurrent selection studies that are being conducted at the Fargo station (Fargo, ND) as an example of maize breeding programs that link
### Table 7.20 Mid-parent (MP) and high-parent (HP) heterosis values for grain yield for intra- and inter-population recurrent selection programs

<table>
<thead>
<tr>
<th>Population hybrids</th>
<th>Avg. heterosis</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS10(FR)C8 × BS11(FR)C8</td>
<td>39.5 34.1</td>
<td>Eyherabide and Hallauer (1991a)</td>
</tr>
<tr>
<td>BSSS(R)C11 × BSCB1(R)C11</td>
<td>76.0 72.4</td>
<td>Keeratinijakal and Lamkey (1993)</td>
</tr>
<tr>
<td>BS21(R)C6 × BS22(R)C6</td>
<td>25.4 22.5</td>
<td>Menz et al. (1999)</td>
</tr>
<tr>
<td>BS21(R)C7 × BS22(R)C7</td>
<td>45.3 43.2</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × CGSS(S1–S2)C5</td>
<td>50.6 43.2</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × CGL(S1–S2)C5</td>
<td>55.8 52.0</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>BS21(R)C7 × NDSAB(MER-FS)C13</td>
<td>31.0 18.2</td>
<td>Carena (unpublished)</td>
</tr>
<tr>
<td>BS22(R)C7 × LEAMING(S)C4</td>
<td>43.0 36.2</td>
<td>Carena (2005)</td>
</tr>
<tr>
<td>LEAMING(S1–S2)C3 × MIDLAND(S1–S2)C3</td>
<td>17.8 7.5</td>
<td>Carena and Hallauer (2001b)</td>
</tr>
<tr>
<td>SynB(S)C6 × CCGPB(RRS)C3</td>
<td>28.9 26.7</td>
<td>Lee et al. (2006)</td>
</tr>
<tr>
<td>SynA(S)C6 × CCGPA(RRS)C3</td>
<td>19.9 16.7</td>
<td>Lee et al. (2006)</td>
</tr>
<tr>
<td>HopeA(RRS)C5 × CGSS(comb)C3</td>
<td>33.0 8.7</td>
<td>Lee et al. (2006)</td>
</tr>
<tr>
<td>Average across population hybrids&lt;sup&gt;a&lt;/sup&gt;</td>
<td>38.9 28.2</td>
<td></td>
</tr>
<tr>
<td>Average across improved hybrids&lt;sup&gt;b&lt;/sup&gt;</td>
<td>18.8 11.1</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Average of selected population hybrids  
<sup>b</sup>Average of 71 variety crosses including 25 improved parent varieties  
Source: Adapted from Carena and Wicks III (2006)

recurrent selection for germplasm improvement with inbred line development in early-maturing regions.

### 7.4 General Effects of Selection

Important information can be obtained from selection experiments besides the direct effect of selection on yield and other traits, e.g., relative changes in heterosis and combining ability effects when the population is crossed with specific or non-specific populations or testers. Lonnquist (1951) selected for high and low yields using the parental population Krug as tester. High- and low-yield synthetics were testcrossed with the single-cross WF9 × M14, and testcross yields were found to be 1.7 and 9.7 q/ha, respectively, below the tester mean. It was concluded that after only one cycle of recurrent selection the original Krug population was separated into two distinct groups with respect to combining ability with an unrelated narrow base tester.
Table 7.21  Populations undergoing recurrent selection for greater grain yield, lodging resistance, drought tolerance, cold tolerance, grain quality, fast dry down, and earlier maturity in the North Dakota State University (NDSU) maize breeding program

<table>
<thead>
<tr>
<th>Selection population</th>
<th>Designation</th>
<th>Tester</th>
<th>Progenies evaluated</th>
<th>Cycles completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Dakota Synthetic AB</td>
<td>NDSAB(MER-FS)(^a)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>NDSAB(M)</td>
<td>—</td>
<td>Half-sib progenies</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>NDSAB(M-DT)</td>
<td>—</td>
<td>Full-sib progenies</td>
<td>16</td>
</tr>
<tr>
<td>North Dakota Synthetic M</td>
<td>NDSM(M-FS)(^a)</td>
<td>—</td>
<td>Half-sibs/full-sibs(^b)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>NDSM(M-CT)</td>
<td>—</td>
<td>Half-sibs/full-sibs(^b)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>NDSS</td>
<td>—</td>
<td>Full-sibs</td>
<td>2</td>
</tr>
<tr>
<td>North Dakota Stiff Stalk</td>
<td>NDSS</td>
<td>—</td>
<td>Full-sibs</td>
<td>2</td>
</tr>
<tr>
<td>Lancaster</td>
<td>NDSS</td>
<td>—</td>
<td>Full-sibs</td>
<td>2</td>
</tr>
<tr>
<td>North Dakota Synthetic CD</td>
<td>NDSCD(M-FS)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>13</td>
</tr>
<tr>
<td>Leaming</td>
<td>LEAMING(S-FS)</td>
<td>—</td>
<td>S(_1)–S(_2)/full-sibs</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>LEAMING(FR)</td>
<td>BS22</td>
<td>Reciprocal full-sibs</td>
<td>1</td>
</tr>
<tr>
<td>Krug Hi I. Synthetic 3c</td>
<td>NDBSK(M-FS)(^b)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>15</td>
</tr>
<tr>
<td>Pioneer Two-Ear Synthetic(^c)</td>
<td>NDBS11(M-FS)(^b)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>17</td>
</tr>
<tr>
<td>Iowa Two-Ear Synthetic (\times) Pioneer Two-Ear Synthetic(^c)</td>
<td>NDBS1011(FR)(^b)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>17</td>
</tr>
<tr>
<td>CYMMIT Highland Comp.(^c)</td>
<td>NDSHLC(M-FS)(^b)</td>
<td>—</td>
<td>Half-sibs/full-sibs</td>
<td>8</td>
</tr>
<tr>
<td>Iowa Early Synthetic No. 1</td>
<td>NDBS21(R-HT)</td>
<td>LH176</td>
<td>Half-sibs</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>NDBS21(FR)</td>
<td>CGSS</td>
<td>Reciprocal full-sibs</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>NDBS21(FR)</td>
<td>NDSAB</td>
<td>Reciprocal full-sibs</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>NDBS21(FR)</td>
<td>CGL</td>
<td>Reciprocal full-sibs</td>
<td>3</td>
</tr>
<tr>
<td>Iowa Early Synthetic No. 2</td>
<td>NDBS22(R-HT)</td>
<td>TR1017</td>
<td>Half-sibs</td>
<td>9</td>
</tr>
<tr>
<td>North Dakota EarlyGEM(^d)</td>
<td>NDBS22(R-CT)</td>
<td>—</td>
<td>Half-sibs</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>EarlyGEM(FS)</td>
<td>—</td>
<td>Full-sibs</td>
<td>2</td>
</tr>
<tr>
<td>CGSS(S)C5 (\times) CGL(S)C5</td>
<td>NDCG(FS)</td>
<td>—</td>
<td>S(_1)–S(_2)/full-sibs</td>
<td>3</td>
</tr>
</tbody>
</table>

\(^a\)Two programs are active with a different number of recombination (Syn 1 vs. Syn 2).
\(^b\)The improved populations BSK(HI)C11 and BS11(FR)C13 were adapted to North Dakota and after three cycles of stratified mass selection for adaptation it was included in the full-sib selection program for genetic improvement. These populations also have ongoing breeding programs from grain quality including oil and protein.
\(^c\)Stratified mass selection was used to adapt tropical and late temperate populations to northern US environments before initiation of selection programs to maximize genetic improvement.
\(^d\)Includes three selection programs where emphasis is (a) grain yield, (b) grain quality (including extractable and fermentable starch,) and (c) earliness (including fast dry down and grain moisture at harvest)
Sprague and Russell (1957) used half-sib family selection in populations Lancaster and Kolkmeier for combining ability with the inbred line Hy. After two cycles, the cross Lancaster × Kolkmeier showed a rate of increase greater than either of the populations themselves. Results were in agreement with those expected on the basis of partial to complete dominance of genes controlling yield. The same conclusion was reported by Sprague et al. (1959) in a further study involving the same populations. Carena and Hallauer (2001a), in a similar study, evaluated response to inbred progeny selection within Leaming and Midland, two open-pollinated varieties identified by Kauffmann et al. (1982) as potential heterotic patterns for the US Corn Belt. After three cycles of selection direct response to selection for grain yield was 9.4 (Leaming) and 11.4% (Midland). Although RRS methods were not used, crosses between the original cultivars and after three cycles of inbred progeny selection were tested. Even though it would be a measure of indirect response, heterosis increased from 4.9% for the C0 × C0 cross to 17.7% for the C3 × C3 cross.

As shown in the previous section, there are more examples of heterotic responses found among diverse populations improved by intra-population recurrent selection methods. Clearly, in these cases, non-additive gene action seems to be important. Lonnquist and Gardner (1961) obtained increases in yield of 5.4 and 4.5% after one cycle of recurrent selection in the varieties Krug and Nubold Reid. The population cross yield increased 13.5% and heterosis increased from 6.0 to 14.6% in one cycle. Crosses between advanced and original populations showed heterosis of 8.3 and 8.7% of mid-parent values for Krug and Nubold Reid, respectively. Results indicated that improved yields of the derived populations were the results of additive gene action present in the parental varieties and genes were in the partial to complete dominance range, thereby resulting in an increase in heterosis. Similar conclusions were reached by Penny et al. (1962) by use of half-sib family selection in Alph and WF9 × B7 for combining ability with the inbred line B14. They observed that crosses between populations in two cycles of selection produced significantly greater yields than the original cross. Selection for poorer yield was effective in the cross between populations in both cycles.

In crosses between three populations within cycles, however, Lonnquist (1963) obtained an average increase in yield that was of the same magnitude as the average increase in yield of the populations themselves. These results indicated that no apparent change occurred in non-additive genetic effects and progressively greater yields in the intercroses were directly related to improvement shown in the parent populations through selection based on additive genetic variation.

Horner et al. (1963) selected for SCA (inbred line F6 as tester) and for GCA (parental population Fla. 767 as tester) during four cycles. Average performance of the selected populations in combination with 11 unrelated testers showed a significant increase in the SCA series but not in the GCA series. In the same manner, SCA with the inbred tester was increased significantly but GCA was not. Horner et al. (1969) evaluated three cycles of recurrent selection based on yield of testcross progeny (inbred tester and population as tester) and of S2 progenies within the Fla. 767 maize population. The improvement in combining ability with 11 unrelated testers was increased significantly but there were no differences among
methods. Horner et al. (1973) compared four cycles involving three methods of recurrent selection and GCA evaluated through testcrosses with unrelated testers was increased significantly by all methods. They showed that the inbred line F6 was nearly twice as effective as a broad base tester or S2 progeny selection for improving frequencies of genes having additive effects. Use of narrow base (inbred line) tester was more effective, indicating that the inbred line was homozygous recessive at many important loci. The inbred tester had greater testcross variances and permitted more successful selection of dominant favorable alleles than a broad base tester. Results obtained by Horner et al. (1976), after selection for SCA with a single-cross tester, showed that selection was as effective for improving GCA as it was for SCA. They speculated about the possibility of changing testers in a recurrent selection program with little loss in accumulated improvement. Such results, as well as those reported by other authors (Sprague et al. 1959; Lonnquist 1961; Russell et al. 1973) have suggested that selection with a specific tester has been effective primarily for additive genetic effects. In a later report, Horner et al. (1989) found that after eight cycles of selection that half-sib family selection was more effective than S2 progeny selection; they attributed the greater gain to overdominant effects with the use of inbred tester F6.

Genter (1973) detected no significant selection effect by either S1 progeny or testcross evaluation in the population VLE. Testcross selection in VCBS was effective in increasing yield in crosses but not in the population itself, whereas S1 selection increased both combining ability and population yield. Genter and Eberhart (1974) evaluated the performance of six original and advanced populations in diallel crosses. Significant improvement in the average cross performance was obtained for VCBS(HT), NHG, PHCB, and PHWI populations but not for BSK and BSSS. However, Vencovsky et al. (1970) evaluated five cycles of mass selection in Paulista Dent (PD) and three cycles in Cateto M. Gerais (CMG) varieties. Heterosis in the cross between the two populations increased in the first cycle but decreased in the third cycle. Greatest heterosis was observed in the cross of PDIII × GMGI, but the greatest yielding cross was PDV × GMGIII with no significant heterosis. Such results showed that improvement in the population cross was mainly at the expense of improvement in the populations themselves through additive genetic effects within populations.

Moll and Robinson (1967) compared full-sib family selection and RRS involving the varieties Jarvis and Indian Chief. After three cycles, full-sib selection resulted in greater population performance than RRS and the population cross increased at a higher rate in full-sib selection. Subsequent reports by Moll and Stuber (1971) and Moll et al. (1978) showed that heterosis of the variety cross from full-sib family selection decreased from the first to the sixth (3.0%) and eighth (12.1%) cycles, but increased after six (8.5%) and eight (11.3%) cycles of RRS (Table 7.22).

Burton et al. (1971) showed that selection in a BSK population through testcross with a double-cross tester and through S1 lines per se both resulted in an increase in mean yield and in GCA. However, S1 selection was more effective. After eight cycles of selection, however, half-sib selection was more effective than inbred progeny selection; no further response to selection was obtained after four cycles of
Table 7.22 Progress per cycle (percent of original mean) through full-sib family (FS) and reciprocal recurrent selection (RRS) in two populations and their cross in four independent evaluations

<table>
<thead>
<tr>
<th>Population</th>
<th>1st (3 cycles)(^a)</th>
<th>2nd (3 cycles)(^b)</th>
<th>3rd (6 cycles)(^c)</th>
<th>4th (8 cycles)(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FS</td>
<td>RRS</td>
<td>FS</td>
<td>RRS</td>
</tr>
<tr>
<td>Jarvis</td>
<td>3.6</td>
<td>4.3</td>
<td>3.5</td>
<td>2.3</td>
</tr>
<tr>
<td>Indian Chief</td>
<td>2.1</td>
<td>1.7</td>
<td>2.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Jarvis × Indian Chief</td>
<td>2.8</td>
<td>0.8</td>
<td>2.5</td>
<td>3.5</td>
</tr>
</tbody>
</table>

\(^a\)Moll and Robinson (1966)
\(^b\)Moll and Stuber (1971)
\(^c\)Moll et al. (1978)

inbred progeny selection (Tanner and Smith 1987). The cross between the advanced populations exhibited heterosis, indicating that the two methods developed populations that differed in gene frequency. Johnson and Salazar (1967) reported that three cycles of mass selection increased average combining ability of a Cuban Yellow Flint variety about 20% with three inbred lines. The same rate of improvement (from 33.6 to 40.5 q/ha) was observed in the variety itself.

Eberhart et al. (1973) reported that for RRS involving BSSS and BSCB1, heterosis increased from 15% in C0 × C0 to 37% in C5 × C5. Similarly, heterosis increased to 34% in the cross between BSCB1 in the fifth cycle of RRS and BSSS in the seventh cycle using the double-cross IA13 as tester. After 11 cycles of RRS, heterosis for the BSSS × BSCB1 population crosses increased from 25.4% for C0 × C0 cross to 76.0% for C11 × C11 cross (Keeratinijakal and Lamkey, 1993). Gevers (1974) observed that heterosis increased from 6% in the original population cross to 10.3 and 11.0% after three cycles of RRS using two procedures for selecting male parents, random selection, and selecting for agronomic traits, respectively.

Walejko and Russell (1977) reported results of five cycles of selection for SCA with the inbred line Hy. Yield changes observed for the population crosses (Lancaster × Kolkmeier) from C0 × C0 to C5 × C5 were at rates of 2.45 ± 0.45 q/ha. Changes in yield of testcrosses with non-specific testers also were similar to those of the Hy testcrosses. A comprehensive evaluation of selection effects suggested that selection was effective for increasing gene frequency of favorable alleles in both populations, and that gene action for yield heterosis was from genes having additive effects with partial to complete dominance.

Changes in mid-parent and variety cross mean yields and in heterosis are summarized in Table 7.23 for the set of intra- and inter-population selection experiments reported in 1981.

Selection for high yield increased mid-parent and variety cross yields in all selection schemes except that reported by Walejko and Russell (1977), where selection was effective for increasing the variety cross yield but not the mid-parent yield. It
### Table 7.23 Changes in mid-parent (MP) and variety cross (VC) mean yield and in heterosis ($H$) after intra- or inter-population recurrent selection

<table>
<thead>
<tr>
<th>Methoda</th>
<th>No. of cycles</th>
<th>MP</th>
<th>VC</th>
<th>$H^b$ (MP)</th>
<th>$H$ (% MP) in C0</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRS</td>
<td>2</td>
<td>5.9</td>
<td>5.8</td>
<td>5.3</td>
<td>3.9</td>
<td>5.3</td>
</tr>
<tr>
<td>HT</td>
<td>2</td>
<td>5.9</td>
<td>2.9</td>
<td>−3.3</td>
<td>3.87</td>
<td>4.2</td>
</tr>
<tr>
<td>HT-high</td>
<td>2</td>
<td>5.9</td>
<td>3.8</td>
<td>−0.3</td>
<td>−2.73</td>
<td>4.8</td>
</tr>
<tr>
<td>HT-low</td>
<td>2</td>
<td>−13.0</td>
<td>−7.2</td>
<td>4.4</td>
<td>−11.21</td>
<td>4.8</td>
</tr>
<tr>
<td>RRS</td>
<td>3</td>
<td>2.9</td>
<td>0.8</td>
<td>−12.0</td>
<td>−6.88</td>
<td>4.2</td>
</tr>
<tr>
<td>FS</td>
<td>3</td>
<td>4.4</td>
<td>3.7</td>
<td>−0.4</td>
<td>−2.13</td>
<td>4.2</td>
</tr>
<tr>
<td>HS</td>
<td>3</td>
<td>3.1</td>
<td>6.1</td>
<td>47.2</td>
<td>2.92</td>
<td>4.2</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>3.9</td>
<td>2.7</td>
<td>−13.0</td>
<td>−1.09</td>
<td>7.4</td>
</tr>
<tr>
<td>HT</td>
<td>5</td>
<td>5.2</td>
<td>7.3</td>
<td>12.6</td>
<td>1.98</td>
<td>4.2</td>
</tr>
<tr>
<td>RRS</td>
<td>5</td>
<td>0.5</td>
<td>4.2</td>
<td>24.8</td>
<td>4.20</td>
<td>4.2</td>
</tr>
<tr>
<td>HT</td>
<td>2</td>
<td>0.2</td>
<td>3.7</td>
<td>29.9</td>
<td>4.07</td>
<td>4.2</td>
</tr>
<tr>
<td>S1</td>
<td>2</td>
<td>4.2</td>
<td>3.4</td>
<td>−2.5</td>
<td>−0.76</td>
<td>4.2</td>
</tr>
<tr>
<td>RRS1</td>
<td>3</td>
<td>7.5</td>
<td>5.8</td>
<td>−7.8</td>
<td>−1.26</td>
<td>5.9</td>
</tr>
<tr>
<td>RRS2</td>
<td>3</td>
<td>3.3</td>
<td>3.3</td>
<td>2.8</td>
<td>0.13</td>
<td>5.9</td>
</tr>
<tr>
<td>HT</td>
<td>5</td>
<td>−0.2</td>
<td>3.6</td>
<td>15.4</td>
<td>5.10</td>
<td>25.4</td>
</tr>
<tr>
<td>RRS-1</td>
<td>1</td>
<td>5.1</td>
<td>7.5</td>
<td>20.6</td>
<td>2.71</td>
<td>18.4</td>
</tr>
<tr>
<td>FS-RRS</td>
<td>3</td>
<td>5.9</td>
<td>3.2</td>
<td>−24.2</td>
<td>−2.49</td>
<td>9.7</td>
</tr>
<tr>
<td>HS</td>
<td>6</td>
<td>4.0</td>
<td>−0.7</td>
<td>−12.9</td>
<td>−5.3</td>
<td>39.2</td>
</tr>
<tr>
<td>RRS</td>
<td>3</td>
<td>2.2</td>
<td>7.1</td>
<td>19.4</td>
<td>6.3</td>
<td>39.2</td>
</tr>
</tbody>
</table>

$^a$RRS: reciprocal recurrent selection; RRS1 and RRS2: males randomly sampled and males selected, respectively; RRS-1: RRS based on testcross of half-sib families; FS: full-sib family selection; FS-RRS: full-sib reciprocal recurrent selection; HT: half-sib (testcross) family selection; HS: half-sib (modified ear-to-row) family selection; $S_1$: $S_1$ family selection

$^b$H = (VC − MP); H (% MP) = 100 H/MP

was concluded that about 50% of the set of studies showed a decrease in intervarietal heterosis. RRS schemes were designed to maximize selection for GCA and SCA effects and therefore increase the variety cross mean through selection for non-additive genetic effects. The limited amount of available experimental data that were available at that moment did not show consistent trends in the effectiveness of selection for increase of heterosis in the variety cross. The apparent failure of RRS to increase heterosis was thought to be attributable to differences in magnitude of effects of genes controlling yield. Since the number of selection cycles for these studies is not high, a logical explanation would be that if genes of larger effects do have mostly additive gene action then selection would change primarily the frequencies of such genes. Therefore, selection for allelic (dominance) or non-allelic (epistasis) interaction would be less effective in the first generations. For five RRS selection programs conducted with more than six cycles of selection, average mid-parent heterosis increased from 7.3% for the original population
crosses to 37.4% for the selected populations (Table 7.24). It seems the original goals of RRS are being attained to improve the population crosses that represent heterotic patterns used in maize breeding (Hallauer and Carena, 2009). Melani and Carena (2005) have proposed alternative heterotic patterns for the northern USA that are currently under full-sib reciprocal recurrent selection where extensive testing is emphasized by pair-cross seed production of S1 progenies instead of using prolific genotypes.

An extensive study by Genter and Eberhart (1974), summarized in Table 7.25, involves 20 variety crosses among original and advanced cycles of intra-population selection. General results do not depart much from those shown in Table 7.23. In only one case mid-parent yield decreased after selection, while 6 out of 20 showed a decrease in the variety cross yield. Heterosis in absolute values decreased in eight of the 20 instances after selection. Heterosis measured in percent of mid-parent decreased in 50% of the crosses. Despite the limited amount of information, results

### Table 7.24 Direct and indirect responses and mid-parent heterosis for reciprocal recurrent selection programs with at least six cycles of selection

<table>
<thead>
<tr>
<th>Populations</th>
<th>References</th>
<th>Cycles of selection no.</th>
<th>Direct %</th>
<th>Indirect %</th>
<th>C0 × C0</th>
<th>Cn × Cn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarvis X Indian Chief</td>
<td>Moll and Hanson (1984)</td>
<td>10</td>
<td>2.7</td>
<td>3.1</td>
<td>6.6</td>
<td>28.9</td>
</tr>
<tr>
<td>BS10 X</td>
<td>Eyherabide and Hallauer (1991)</td>
<td>8</td>
<td>7.5</td>
<td>3.0</td>
<td>2.5</td>
<td>39.6</td>
</tr>
<tr>
<td>BS11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>BSSS X</td>
<td>Keeratinijakal and Lamkey (1993)</td>
<td>11</td>
<td>7.0</td>
<td>2.0</td>
<td>25.4</td>
<td>76.0</td>
</tr>
<tr>
<td>BSCB1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>BS21 X</td>
<td>Menz et al. (1999)a</td>
<td>6</td>
<td>4.4</td>
<td>−0.2</td>
<td>1.0</td>
<td>25.4</td>
</tr>
<tr>
<td>BS22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>BS21 X</td>
<td>Menz et al. (1999)b</td>
<td>6</td>
<td>1.6</td>
<td>−5.9</td>
<td>1.0</td>
<td>17.2</td>
</tr>
<tr>
<td>BS22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−0.5</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>4.6</td>
<td>0.3</td>
<td>7.3</td>
<td>37.4</td>
<td></td>
</tr>
</tbody>
</table>

*aPopulations used as testers
*bInbred lines used as testers: A632 for BS21 and H99 for BS22
Source: Adapted from Hallauer and Carena (2009)
suggest that changes in heterosis after intra-population selection are largely a matter of chance. However, crosses between genetically distant populations improved by intra-population selection programs also can identify population hybrids with significantly greater heterosis expression (see Table 7.19). Therefore, probabilities of finding good heterotic combinations in intra-population selection programs increase with good choice of germplasm and how improved populations are.

Theory and empirical results have shown that selection directly affects the variability within populations. The extent of changes in variability depends on several factors, e.g., selection intensity, initial amount of variability (actual and potential), linkage disequilibrium, and rate of recombination. Environmental factors may affect the recombination rate, so the gradual liberation of potential variability also depends on environmental conditions to some extent (Mock, 1973).

Table 7.25 Changes in mid-parent (MP) and variety cross (VC) mean yield and in heterosis ($H$) after intra-population recurrent selection

<table>
<thead>
<tr>
<th>Populations$^a$</th>
<th>Methods$^b$</th>
<th>No. of cycles</th>
<th>MP in % of C0</th>
<th>VC in % of C0</th>
<th>$H^b$ in % MP</th>
<th>$H^c$ (% MP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A B</td>
<td>A B</td>
<td>A B</td>
<td>MP</td>
<td>VC</td>
<td>$H$</td>
<td>$H^b$</td>
</tr>
<tr>
<td>VCBS BSK HT S$_1$</td>
<td>3 4</td>
<td>16.2</td>
<td>18.4</td>
<td>30.6</td>
<td>1.5</td>
<td>13.0</td>
</tr>
<tr>
<td>VCBS BSSS HT HT</td>
<td>3 7</td>
<td>8.9</td>
<td>18.0</td>
<td>73.9</td>
<td>9.7</td>
<td>16.3</td>
</tr>
<tr>
<td>VCBS NHG HT M</td>
<td>3 12</td>
<td>17.0</td>
<td>8.1</td>
<td>-44.5</td>
<td>-8.9</td>
<td>17.0</td>
</tr>
<tr>
<td>VCBS PHCB HT M</td>
<td>3 9</td>
<td>14.9</td>
<td>23.9</td>
<td>98.3</td>
<td>8.8</td>
<td>12.1</td>
</tr>
<tr>
<td>VCBS PHWI HT M</td>
<td>3 9</td>
<td>8.4</td>
<td>17.8</td>
<td>117.8</td>
<td>9.4</td>
<td>9.4</td>
</tr>
<tr>
<td>VCBS BSK S$_1$ S$_1$</td>
<td>4 4</td>
<td>16.0</td>
<td>18.2</td>
<td>30.6</td>
<td>1.6</td>
<td>13.0</td>
</tr>
<tr>
<td>VCBS BSSS S$_1$ HT</td>
<td>4 7</td>
<td>8.7</td>
<td>19.3</td>
<td>84.8</td>
<td>11.4</td>
<td>16.3</td>
</tr>
<tr>
<td>VCBS NHG S$_1$ M</td>
<td>4 12</td>
<td>16.8</td>
<td>15.9</td>
<td>11.0</td>
<td>-0.9</td>
<td>17.0</td>
</tr>
<tr>
<td>VCBS PHCB S$_1$ M</td>
<td>4 9</td>
<td>14.7</td>
<td>23.9</td>
<td>100.0</td>
<td>9.0</td>
<td>12.1</td>
</tr>
<tr>
<td>VCBS PHWI S$_1$ M</td>
<td>4 9</td>
<td>8.2</td>
<td>29.3</td>
<td>149.5</td>
<td>12.2</td>
<td>9.4</td>
</tr>
<tr>
<td>BSK BSSS S$_1$ HT</td>
<td>4 7</td>
<td>2.8</td>
<td>-14.6</td>
<td>-89.0</td>
<td>-20.9</td>
<td>23.4</td>
</tr>
<tr>
<td>BSK NHG S$_1$ M</td>
<td>4 12</td>
<td>11.2</td>
<td>-0.2</td>
<td>-75.7</td>
<td>-11.8</td>
<td>15.1</td>
</tr>
<tr>
<td>BSK PHCB S$_1$ M</td>
<td>4 9</td>
<td>8.9</td>
<td>-6.5</td>
<td>-72.9</td>
<td>-17.4</td>
<td>23.2</td>
</tr>
<tr>
<td>BSK PHWI S$_1$ M</td>
<td>4 9</td>
<td>2.7</td>
<td>-5.4</td>
<td>-77.6</td>
<td>-8.7</td>
<td>11.2</td>
</tr>
<tr>
<td>BSSS NHG HT M</td>
<td>7 12</td>
<td>4.3</td>
<td>-4.7</td>
<td>-37.4</td>
<td>-11.0</td>
<td>27.6</td>
</tr>
<tr>
<td>BSSS PHCB HT M</td>
<td>7 9</td>
<td>2.0</td>
<td>1.3</td>
<td>-2.1</td>
<td>-0.8</td>
<td>19.2</td>
</tr>
<tr>
<td>BSSS PHWI HT M</td>
<td>7 9</td>
<td>-3.4</td>
<td>3.3</td>
<td>45.9</td>
<td>-8.0</td>
<td>15.7</td>
</tr>
<tr>
<td>NHG PHCB M M</td>
<td>12 9</td>
<td>10.1</td>
<td>12.7</td>
<td>29.8</td>
<td>2.7</td>
<td>14.8</td>
</tr>
<tr>
<td>NHG PHWI M M</td>
<td>12 9</td>
<td>4.2</td>
<td>8.1</td>
<td>85.7</td>
<td>4.0</td>
<td>5.1</td>
</tr>
<tr>
<td>PHCB PHWI M M</td>
<td>9 9</td>
<td>1.9</td>
<td>-1.1</td>
<td>-24.5</td>
<td>-3.4</td>
<td>13.3</td>
</tr>
</tbody>
</table>

$^a$VCBS: Virginia Corn Belt Southern Synthetic; BSK: Krug Hi I Synthetic 3; BSSS: Stiff Stalk Synthetic; NHG: Nebraska Hays Golden; PHCB: Pioneer Hi-Bred Corn Belt Synthetic; PHWI: Pioneer Hi-Bred West Indian Synthetic
$^b$HT: half-sib (testcross) family selection; S$_1$: S$_1$ family selection; M: mass selection
$^c$H = (VC - MP); H (% MP) = 100 H/MP

Source: Adapted from Genter and Eberhart (1974)
One of the earliest reports on changes in variability by selection was after selection for protein and oil in Burr’s White in the period of 1896–1924 (Winter, 1929). The measure of variability through three different methods (Weinberg’s formula, standard deviation, and extra-modal coefficient) showed an increasing trend for high protein and high oil strains but a decreasing trend in low protein and low oil strains. Variability measured through the coefficient of variation showed a decreasing trend in high oil and high protein strains because the mean of the traits increased in both cases. In the same way the coefficient of variation showed an increasing trend in low protein and low oil strains because the mean decreased in both cases during the selection period. Leng (1962) demonstrated the existence of variability after 48 generations as indicated by effective reverse selection in all four strains. Dudley and Lambert (1969) estimated the genetic variability in five strains of the Illinois selection program. Estimates of variance among half-sib families were significant in all cases but were greater in the strains selected for decreasing traits (low oil and low protein). A direct comparison of different estimates was not available because of non-homogeneity of error variances. For this reason a rough measure of heritability on a single-plot basis was presented for comparisons (Table 7.26).

<table>
<thead>
<tr>
<th>Population</th>
<th>Oil</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\sigma}_g^2$</td>
<td>$\hat{\sigma}_g^2/\hat{\sigma}_p^2$ (%)</td>
</tr>
<tr>
<td>Illinois high oil (IHO)</td>
<td>0.0724</td>
<td>12.5</td>
</tr>
<tr>
<td>Illinois low oil (ILO)</td>
<td>0.0008</td>
<td>16.7</td>
</tr>
<tr>
<td>Illinois high protein (IHP)</td>
<td>0.0438</td>
<td>39.8</td>
</tr>
<tr>
<td>Illinois low protein (ILP)</td>
<td>0.0090</td>
<td>15.5</td>
</tr>
<tr>
<td>Illinois high protein (HN)</td>
<td>0.0210</td>
<td>30.0</td>
</tr>
</tbody>
</table>

\[\hat{\sigma}_p^2\] is phenotypic variance on a plot basis: $\hat{\sigma}_g^2 + \hat{\sigma}_d^2 + \hat{\sigma}_y^2$.

Source: Adapted from Dudley and Lambert (1969)

Heritability for non-selected characters was expected to be greater than for selected characters in each population, because changes in variability without selection would be attributable only to inbreeding whereas selected characters would have reduced variability because of both inbreeding and selection. This seemed to have occurred in populations ILO, IHP, and IHP(HN). Heritability for oil content in ILO was slightly greater than in IHO. Variability in ILO, however, was due to the variation in percent of germless kernels; such variability would have limited usefulness for selection toward less oil content (Dudley and Lambert, 2004).

Robinson and Comstock (1955) estimated additive and dominance variances in several hybrid populations and open-pollinated varieties in different cycles of selection. As a result of selection a decrease in the additive genetic variance estimate was observed in the populations of Jarvis, NC34 × NC45, and CI21 × NC7, but a small increase occurred in the Weekly variety. Subsequently, Moll and Robinson (1966)
reported estimates of additive genetic variance in several cycles of intra-population selection and of variance among testcrosses in RRS. Results are summarized in Table 7.27.

Table 7.27  Additive genetic variance ($\hat{\sigma}_A^2$) and testcross progeny ($\hat{\sigma}_g^2$) estimates in several cycles of selection in intra- and inter-population improvement, respectively

<table>
<thead>
<tr>
<th>Population</th>
<th>Cycles of selection</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Additive genetic variance$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI121 × NC7</td>
<td></td>
<td>24±3</td>
<td>20±22</td>
<td>20±14</td>
<td>41±21</td>
<td>18±13</td>
<td>40±21</td>
<td>16±14</td>
</tr>
<tr>
<td>Jarvis</td>
<td></td>
<td>30±4</td>
<td>12±13</td>
<td>16±19</td>
<td>42±18</td>
<td>38±20</td>
<td>53±22</td>
<td>—</td>
</tr>
<tr>
<td>Indian Chief</td>
<td></td>
<td>17±4</td>
<td>1±28</td>
<td>31±17</td>
<td>30±19</td>
<td>44±23</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Variance among testcross progenies$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jarvis × Indian Chief</td>
<td></td>
<td>7±2</td>
<td>18±8</td>
<td>16±8</td>
<td>19±6</td>
<td>9±4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Indian Chief × Jarvis</td>
<td></td>
<td>9±2</td>
<td>22±9</td>
<td>10±6</td>
<td>28±8</td>
<td>23±8</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

$^a \times 10^{-4}$, for yield in pounds per plant
Source: Adapted from Moll and Robinson (1966)

Intra-population estimates of the quantity $\hat{\sigma}_A^2 + \sigma_{AE}^2$ (additive + additive–environment interaction variance) for successive selection cycles seemed to be distributed around the more precise estimate for the original population and revealed no trend associated with selection. Variances among testcross progenies after the first cycle of inter-population selection were greater than estimates in the original population cross. Although no obvious trend was observed in the magnitude of the sequential estimates after the first cycle, data suggested that genetic variance increased after the initial cycle of selection.

Silva and Lonnquist (1968) detected a decrease in genetic variances for yield and days to flower after one cycle of S1 family and of full-sib family selection. The two methods of selection apparently produced different changes in magnitude of genetic variances. Hallauer (1970) estimated additive genetic variance in the original two populations and after four cycles of RRS. Selection was not effective in increasing yield of Corn Borer Synthetic No. 1 and additive genetic variance estimates showed no difference between cycles:

$$C0 (\hat{\sigma}_A^2 = 143 \pm 35) \text{ vs. } C4 (\hat{\sigma}_A^2 = 133 \pm 35)$$

On the other hand, selection was effective in increasing yield of the Iowa Stiff Stalk Synthetic population and estimates of additive genetic variance showed a decreasing trend from:

$$C0 (\hat{\sigma}_A^2 = 184 \pm 48) \text{ to } C4 (\hat{\sigma}_A^2 = 130 \pm 35)$$

Also, the population cross showed an increase in yield and a decrease in $\hat{\sigma}_A^2$ estimates from:
C0 ($\hat{\sigma}^2_A = 216 \pm 46$) to C4 ($\hat{\sigma}^2_A = 96 \pm 30$)

Summaries presented in Chapter 5, however, show no evidence of change in genetic variability with selection.

Hallauer (1971) studied changes in several traits after four cycles of RRS for yield in maize. Additive genetic variance decreased and dominance variance increased for kernel depth in all three populations. For all other traits estimates of additive genetic variance were significant but showed small differences among the three pairs of populations. The change in maturity for C0 and C4 seemed to have influenced estimates of additive genetic variance for silking date and plant and ear heights. Betran and Hallauer (1996a) used the design II mating scheme to determine the inter-population genetic variability after nine cycles of RRS in BSSS and BSCB1. The additive genetic component of variance was the most important component of variance for the 10 traits measured. Grain yield was the more important trait included in selection, and the additive genetic variance increased with selection for greater grain yield and the additive by environment interaction component decreased from the C0 $\times$ C0 to the C9 $\times$ C9. Except for grain yield, the estimates of dominance variance component decreased after nine cycles of RRS for the other nine traits.

Genetic variance estimates in the mass selection program in the Hays Golden open-pollinated variety have shown that variability was consistently reduced in the selected populations (Gardner, 1969b; Harris et al., 1972). Gardner (1976) summarized results relative to effect of selection on the variability of Hays Golden (control and irradiated) variety, as shown in Table 7.28.

<table>
<thead>
<tr>
<th>Population</th>
<th>6</th>
<th>10</th>
<th>15</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hays Golden</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>HG – control</td>
<td>136</td>
<td>30</td>
<td>60</td>
<td>40</td>
<td>22</td>
<td>30</td>
<td>77</td>
</tr>
<tr>
<td>HG – irradiated</td>
<td>263</td>
<td>89</td>
<td>49</td>
<td>67</td>
<td>47</td>
<td>14</td>
<td>57</td>
</tr>
</tbody>
</table>

*Source: Adapted from Gardner (1976)*

In a number of selection experiments the genetic coefficient of variation has been used to measure the relative amount of genetic variability among entries (half-sibs, full-sibs, selfed, or testcross families). A summary of results of several selection experiments relative to the genetic coefficient of variation is presented in Table 7.29 for half-sib and S1 family selection (intra-population) and in Table 7.30 for inter-population improvement and recurrent selection with unrelated testers.
Table 7.29  Genetic coefficient of variation estimates within populations under half-sib and $S_1$ family selection

| Cycle | Modified ear-to-row selection | | | | | |
|-------|-------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|       | Paulista Dent$^a$ | Hays Golden$^b$ | Pinamex$^c$ | Dent Composite$^d$ | $S_1$ selection BSK(S) | $S_2$ selection BS13 |
| 0     | 15.3 | 11.3 | 10.6 | 7.4 | 16.2 | 14.6 |
| 1     | 9.3  | 4.2  | 6.1  | 7.3  | 15.7 | 22.3 |
| 2     | 9.1  | 4.1  | 5.0  | —    | 10.8 | 12.9 |
| 3     | 7.1  | 5.0  | 3.4  | —    | 15.4 | —    |
| 4     | —    | 4.8  | 6.5  | —    | 8.8  | —    |
| 5     | —    | —    | —    | —    | 13.4 | —    |
| 6     | —    | —    | —    | —    | 20.6 | —    |
| 7     | —    | —    | —    | —    | 39.7 | —    |

$^a$Paterniani (1967)
$^b$Webel and Lonnquist (1967)
$^c$Paterniani (1969)
$^d$Miranda et al. (1972); Lima et al. (1974)

Table 7.30  Genetic coefficient of variation estimates in the population cross under reciprocal recurrent selection and half-sib selection with unrelated testers

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Half-sib selection with tester BSSS$^{(1)}$</th>
<th>BS12(HI)$^{(2)}$</th>
<th>BSK(HI)$^{(3)}$</th>
<th>Reciprocal RS BSCB1$^{(4)}$</th>
<th>BSSS$^{(5)}$</th>
<th>Modified RRS-2$^{a}$ Flint Composite$^{(6)}$</th>
<th>Dent Composite$^{(7)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10.2</td>
<td>11.6</td>
<td>4.0</td>
<td>10.8</td>
<td>7.9</td>
<td>11.7</td>
<td>9.5</td>
</tr>
<tr>
<td>1</td>
<td>3.8</td>
<td>7.8</td>
<td>5.2</td>
<td>7.7</td>
<td>6.2</td>
<td>5.7</td>
<td>6.4</td>
</tr>
<tr>
<td>2</td>
<td>4.8</td>
<td>2.8</td>
<td>2.4</td>
<td>3.2</td>
<td>3.7</td>
<td>12.3</td>
<td>7.7</td>
</tr>
<tr>
<td>3</td>
<td>—</td>
<td>4.5</td>
<td>3.4</td>
<td>3.5</td>
<td>2.8</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>2.6</td>
<td>2.4</td>
<td>3.4</td>
<td>3.6</td>
<td>2.1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>3.8</td>
<td>5.0</td>
<td>5.0</td>
<td>4.9</td>
<td>5.3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>4.5</td>
<td>9.7</td>
<td>7.7</td>
<td>8.0</td>
<td>5.5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>—</td>
<td>—</td>
<td>8.5</td>
<td>5.6</td>
<td>4.6</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3.9</td>
<td>8.6</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

$^a$Paterniani (1971, 1974a, 1974b); testers: (1) IA13; (2) B14; (3) IA4652, B14, B73; (4) BSSS; (5) BSCB1; (6) Dent Composite; (7) Flint Composite

A common feature of Tables 7.29 and 7.30 is that the genetic coefficients of variation generally decrease sharply after the first cycle of selection and remain either unchanged or change very little in subsequent cycles. Selection, of course, is expected to be more effective in the first cycle because a greater amount of genetic variability is available. In subsequent cycles an increase in precision of experiments is required to make selection as effective as in the first cycle. Effective selection
increases the mean of the selected trait, which leads to a decrease in the genetic coefficient of variation even if genetic variability remains constant. Because estimates of genetic coefficients of variation are obtained in different years, effects of environment and of genotype–environment interaction may have some influence on the magnitude of genetic variability. In general, empirical results to date have shown that selection seems to have decreased genetic variability, although in most instances only a limited number of cycles of selection have been completed (see Chapter 5). It depends to a great extent on the number of families saved for reproduction in each cycle.

Reeder et al. (1987), for example, examined what genetic changes had occurred after six cycles of RRS, based on full-sib family selection, in BS10 and BS11. Grain yield of full-sib families increased 6.3 and 5.7%, respectively, in BS10 and BS11, whereas grain yield of $S_1$ progenies of the same population increased 11.6 and 26.3%, respectively. The changes in the estimates of genetic variances within BS10 and BS11 populations suggested a trend that genetic variability was reduced after six cycles of selection but most of the changes were small and were not statistically significant. The possible responses realized by full-sib family and inbred progeny selection were the stimulus that prompted Moreno-Gonzalez and Hallauer (1982) of the possibilities of basing selection on two different types of progenies.

A very important feature of recurrent selection methods is their effect on the potential of improved populations as sources of inbred lines for hybrid development. Theoretical and empirical studies have suggested that the increase in frequencies of favorable alleles also increases the opportunity for the development of outstanding hybrids. Horner et al. (1973) suggested that commercial maize hybrids could be developed rapidly from a recurrent selection program using a seed parent as tester already in commercial use. Using this procedure they obtained a hybrid (Florida 200A) that was released for commercial production. B73, however, is the key example on how germplasm improvement with recurrent selection methods can provide commercial inbred lines and recycled versions of it.

Hallauer (1973) reported that full-sib progenies ($S_0 \times S_0$) after one cycle of reciprocal full-sib selection were at a higher yield level than crosses from the unimproved population. The two parent populations, Iowa Two-Ear Synthetic and Pioneer Two-Ear Composite, had low frequencies of prolific plants at normal plant densities. After one cycle of selection the increase in yield was 14.8 and 18.7%, respectively. Also, 46% of the full-sib families after selection exceeded the mean of the checks and 16% were one or more standard deviations above the mean of the checks. Only 1% of the full-sib families from crosses between the original populations exceeded the mean of the same checks.

Suwantaradon and Eberhart (1974) observed that hybrids developed from two improved populations had significantly higher yield than the cross between the parent varieties. The parent varieties BSK and BSSS were improved through five cycles of $S_1$ family selection and RRS with BSCB1 as tester, respectively. The cross between improved varieties yielded at least 90% as much as the best hybrid checks and the best hybrid developed from them yielded 18% more than the variety cross. Betran and Hallauer (1996b) also compared the means of the hybrids produced from
crosses of the C0 and C9 populations. The single crosses produced after nine cycles of RRS had significantly greater average yields (2.67 t/ha or 54.5%) than the average of the C0 × C0 single crosses. RRS was more effective than the combined response (2.09 t/ha or 42.8%) of the half-sib and inbred progeny recurrent selections conducted in BS13 for increased grain yield. RRS also was more effective for increasing root and stalk strength and reducing ear height and days to flower.

Gardner (1972) developed S2 lines from three sources: (1) the parent variety Hays Golden (HG) population (control) after 12 cycles of mass selection for yield (C12); (2) from the irradiated population after 13 cycles of mass selection for yield (I13); and (3) from the prolific (P7) population after seven cycles of mass selection for number of ears per plant. Inbred lines from the selected populations were superior in yield to lines from the parent variety when crossed with oh43 (inbred line) and the single cross of related lines or often known as sister lines (N7A × N7B). Lines from C12, I13, and P7 exceeded lines from HG by 11.4, 10.0, and 10.6%, respectively, when crossed with oh43 and by 10.9, 8.0, and 7.7% when crossed with N7A × N7B. As a result 10 hybrids including these lines exceeded the best check hybrid in each experiment.

Harris et al. (1972) evaluated S1 lines themselves and in testcross performance after nine cycles of mass selection for yield in Hays Golden. It was shown that the S1 lines from the selected populations were superior in mean yield to lines of the parent variety either as S1 lines themselves or in testcrosses. It was concluded that selection eliminated radiation-induced deleterious mutants from the irradiated population while increasing frequencies of favorable yield genes common to both selected populations. This apparently produced similar germplasm reservoirs more suitable than the parent variety for the development of superior inbred lines. In addition, Martin and Gardner (1976) compared single-cross, three-way, and double-cross hybrids from inbred lines of Hays Golden and from the control and irradiated derived populations selected after nine cycles of mass selection. Hybrids developed from the mass selected control and irradiated populations were 9.3 and 7.4% higher yielding and 7.4 and 17.6% more prolific, respectively, than hybrids developed from the original Hays Golden variety.

7.5 Factors Affecting Efficiency of Selection

Different factors affecting efficiency of selection were discussed in detail by Eberhart (1970). He developed the prediction formula that has general application to the selection methods used in maize breeding and showed how the variables in the prediction formula can be manipulated to increase efficiency of selection. Prediction equations for different selection methods are given in Chapter 6 and were discussed to some extent.

An example to illustrate the prediction formula use is given in Tables 7.31 and 7.32. Three traits were measured for 144 S1 progenies of Iowa Stiff Stalk Synthetic across three replications in one experiment. Ten plants within each plot
### Table 7.31 Analysis of variance and covariance for 144 BSSS S<sub>1</sub> lines for first-brood European corn borer, stalk rind puncture, and stalk rot rating compared in one experiment

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>Rind puncture</th>
<th>Corn borer</th>
<th>Stalk rot</th>
<th>Mean cross-products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>2</td>
<td>18.86</td>
<td>3.02</td>
<td>0.48</td>
<td>4.49</td>
</tr>
<tr>
<td>Entries</td>
<td>143</td>
<td>14.62</td>
<td>5.67</td>
<td>2.19</td>
<td>−1.57</td>
</tr>
<tr>
<td>Error</td>
<td>286</td>
<td>2.10</td>
<td>2.55</td>
<td>0.60</td>
<td>1.05</td>
</tr>
<tr>
<td>Total</td>
<td>431</td>
<td></td>
<td></td>
<td></td>
<td>−0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>Rind puncture</th>
<th>Corn borer</th>
<th>Stalk rot</th>
<th>Corn borer × Rind puncture</th>
<th>Corn borer × Stalk rot</th>
<th>Corn borer × Stalk rot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>2</td>
<td>18.86</td>
<td>3.02</td>
<td>0.48</td>
<td>4.49</td>
<td>−1.57</td>
<td>−1.20</td>
</tr>
<tr>
<td>Entries</td>
<td>143</td>
<td>14.62</td>
<td>5.67</td>
<td>2.19</td>
<td>1.05</td>
<td>−3.97</td>
<td>0.37</td>
</tr>
<tr>
<td>Error</td>
<td>286</td>
<td>2.10</td>
<td>2.55</td>
<td>0.60</td>
<td>0.13</td>
<td>−0.25</td>
<td>−0.03</td>
</tr>
<tr>
<td>Total</td>
<td>431</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \hat{\sigma}^2 = 2.10 \pm 0.17 \]
\[ \hat{\sigma}^2_g = 4.17 \pm 0.17 \]
\[ \hat{\sigma}^2_g = 85.6 \]
\[ \hat{\sigma}^2_g = 10.8 \]
\[ \hat{\sigma}^2_g = 13.4 \]

\[ h^2 (%) = \frac{\hat{\sigma}^2_g}{(\hat{\sigma}^2 + \hat{\sigma}^2_g)} \times 100 \]

\[ \bar{x} = 13.4 \]

\[ r_g = 0.15 \]

\[ r_p = 0.12 \]

\[ r = 0.12 \]

\[ CV(\%) = 10.8 \]

\[ CV(\%) = 5.7 \]

\[ CV(\%) = 2.7 \]

\[ CV(\%) = 28.0 \]

\[ CV(\%) = 28.6 \]
were inoculated for European corn borer (*O. nubilalis* Hübn) and for a stalk rot organism (*D. zea* Pass.). Approximately 3 weeks after inoculation a mean plot rating was made for resistance to first brood European corn borer leaf feeding. Also, about 3 weeks after inoculation with *D. zea*, 10 plants within each plot were measured to determine the amount of pressure required to penetrate the rind of the stalk; later the same 10 plants were sliced with a knife to determine the level of *D. zea* infection. The rind puncture and *D. zea* ratings were made between the second and third internodes above ground level. All analyses were made on plot means. The analyses of variance for each trait and covariances between the three traits are given in Table 7.31. Differences among entries were significant for the three traits, and a negative genetic correlation existed between rind puncture and level of *D. zea* infection. Twenty *S*₁ progenies were selected for recombination to form the next cycle of selection based on measurements of the three traits. For illustration we will assume selection was for each trait. Our problem is to determine if effective selection can be made on three replications in one environment or if more or fewer replications should be used. Expected gains for each trait for different numbers of replications are shown in Table 7.31. Because data were collected in only one environment, relative efficiency of increasing replications for each trait is directly related to level of heritability of the trait. The largest heritability (85.6%) was from rind puncture.
reading. Therefore, the effect of increasing replications was less than for corn borer reading, which had the lowest heritability estimate (55.0%). Expected genetic gain for rind puncture reading increased only 6% by increasing replications from two to five, whereas an 18% increase in genetic gain is predicted for corn borer rating. It seems that two to three replications would be adequate for each trait, but the effect of increasing replications would be greater for corn borer rating. Doubling the number of replications (which doubles land area and labor) increases expected genetic gain for rind puncture reading by only 6%. Because data were collected in only one environment, we have biased estimates of heritability too. It is not possible from these data to determine how effective two replications in two different environments would be for increasing genetic gain. If the genotype by environment interaction component is relatively large compared to the genotypic component, distribution of replications in different environments would be needed.

Rogers et al. (1977) and Russell et al. (1978) have computed expected gains from selection for corn rootworms (Diabrotica spp.) and second-brood European corn borer, respectively, from data collected from different environments. Both studies included S1 progenies that were evaluated for resistance to the particular maize pest under study. Rootworm resistance was dependent on natural infestation, whereas artificial infestation techniques were used for the second-brood European corn borer resistance. Rogers et al. (1977) studied four synthetic varieties undergoing recurrent selection for rootworm tolerance. Four root traits (root lodging, root damage, root size, and secondary roots) were studied in experiments conducted in two environments in each of 2 years, each environment including two replications. Except for root damage, most genotypic and genotype by environment components of variance were significantly different from zero. Expected gains from selection for rootworm tolerance were determined for selection based on each of the four root traits and selection indexes including the four root traits. Indirect selection using root size, secondary roots, and root damage was not as effective in reducing root lodging as direct selection for root lodging itself. The only situation in which direct selection would not be as effective in reducing root lodging would be in environments in which conditions that promote root lodging (e.g., effects of wind and rain) do not occur. Index selection would also be more effective in non-root lodging environments. Because root lodging resistance seemed to be the best trait for selecting for rootworm tolerance, Rogers et al. (1977) calculated expected genetic gain for different combinations for number of replications and environments. Figure 7.4 shows expected gain for three populations. (The fourth population, BSLR, was not included because root lodging occurred in only one environment; hence, no estimate of genotype by environment interaction was available.

Expected progress among the three populations differed because of the magnitudes of genotype (smallest for BSSS) and genotype by environment (smallest for BS1) components of variance. Heritability estimates of root lodging were 45.2, 41.8, and 84.6% for BSSS, BSER, and BS1, respectively. Substantial gains can be expected in all instances by increasing replications from one to two, but additional locations have a greater impact than additional replications as seen in the heritability formula from previous chapters. The value of adding more than two replications
decreases rapidly. For these three populations, it seems that an effective compromise of two replications at three to four environments would be most desirable. For a trait such as root lodging that is very dependent on environmental conditions for its expression, response to selection would be enhanced by increasing the number of locations as permitted by resources available.

In addition to replications and environments, Russell et al. (1978) had individual plant data. They included 100 $S_1$ progenies from each of two synthetic populations BS9C1 and BS16 in their study. BS9C1 had undergone one cycle of recurrent selection for resistance to European corn borer, whereas BS16 had not undergone any previous selection for resistance. The $S_1$ progenies from BS9C1 were evaluated at one location in 1975 and 1976, whereas the $S_1$ progenies from BS16 were evaluated only in 1976. Components of variance estimated from the analyses of variance are given in Table 7.33.

Estimates of components of variance were similar, but because the $S_1$ progenies of BS9C1 were evaluated in each of the 2 years, the BS9C1 estimates are used to illustrate the expected genetic responses from selection.

Selection for resistance to infestation by the second brood of the European corn borer usually is determined by family means obtained by averaging over plots, replications, and years, where data are obtained on 10 plants in each plot for each of
Table 7.33  Variance component estimates obtained for S1 progenies of BS9C1 and BS16 evaluated for second-brood European corn borer resistance (Russell et al., 1978)

<table>
<thead>
<tr>
<th>Population</th>
<th>$\hat{\sigma}^2_w$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{\sigma}^2_{ge}$</th>
<th>$\hat{\sigma}^2_g$</th>
<th>$\hat{h}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS9C1</td>
<td>120.3</td>
<td>19.1</td>
<td>15.1</td>
<td>38.3</td>
<td>78.1</td>
</tr>
<tr>
<td>BS16CO</td>
<td>117.8</td>
<td>14.5</td>
<td>—</td>
<td>31.6</td>
<td>86.7</td>
</tr>
</tbody>
</table>

$\hat{\sigma}^2_w$, $\hat{\sigma}^2$, $\hat{\sigma}^2_{ge}$, and $\hat{\sigma}^2_g$ refer to the within-plot variation, experimental error, genotype by environment interaction, and genotypic components of variance, respectively; $\hat{\sigma}^2_g$ is the variation among S1 progenies and $\hat{\sigma}^2_g = \hat{\sigma}^2_A$ for $p = q = 0.5$ or no dominance.

Heritability calculated as: $\hat{\sigma}^2_g / (\hat{\sigma}^2 / 6 + \hat{\sigma}^2_{ge} / 2 + \hat{\sigma}^2_g) \times 100$

three replications for each year. Figure 7.5 illustrates the effects of number of plants and replications on expected genetic gain per cycle.

Fig. 7.5  Expected genetic gain per cycle for second-brood European corn borer resistance for different combinations of plants per plot and replications in BS9C1 with S1 inbred progeny recurrent selection.

Increase in expected genetic gain is minimal when number of plants is increased beyond 10 in all instances. Comparatively large reductions in genetic gain are expected when four or fewer replications with five or fewer plants per plot are used. For example, with three replications, a decrease from 10 to 5 plants per plot...
decreases expected gain only 8%, whereas a decrease from 5 to 1 plant per plot decreases expected gain 35%. As the number of replications is increased, effect of plot size diminishes. Relation of expected genetic gain to number of replications is similar to that for number of plants per plot. Increasing replications from one to two is slightly greater than increasing replications from four to eight. Also, an increase in plot size diminishes the effects of added replications. Figure 7.4 shows that reducing plot size by 50% to five plants and increasing replications by 33% to four do not alter expected gain, but the expensive and time-consuming procedures of infesting second-brood egg masses and counting stalk cavities are reduced by nearly 30%.

Information on second-brood European corn borer resistance is not available until after flowering. If off-season nurseries are available and 1 year’s data are sufficient to identify the superior S₁ progenies, one cycle of S₁ progeny evaluation can be completed in 2 years. If additional years of data are required, it would increase the duration of each cycle from 2 to 3 years. Because the BS9C1 progenies were evaluated in 2 years, the relation of numbers of replications and years to expected gain was examined on a per-cycle and per-year basis. The effect of increasing the years of evaluation is positive when expected gain is expressed on a per-cycle basis (Fig. 7.6).

![Graph showing expected genetic gain relative to numbers of replications and years on a per-cycle and per-year basis for S₁ selection (Russell et al., 1978)](image-url)

**Fig. 7.6** Expected genetic gain relative to numbers of replications and years on a per-cycle and per-year basis for S₁ selection (Russell et al., 1978)
Results of Fig. 7.6 indicate that reduction in phenotypic variance from increasing number of years of evaluation is not great enough to compensate for increased duration of a cycle of selection. It seems, therefore, that the rate of genetic gain from use of $S_1$ progenies is maximized by evaluating five plants per plot in four replications in 1 year. If genotype by environment interaction is important in the classification of $S_1$ progenies for second-brood corn borer resistance, use of more than one location within each year would be preferable if facilities permit.

Genetic gain expected for different cyclical selection methods for Iowa Stiff Stalk Synthetic (BSSS) is given to illustrate how genetic gain (per cycle and per year) varies among selection methods. Genetic gain is given for situations similar to those of the US Corn Belt; only one season is available for testing but winter nursery facilities can be used for recombination and development of progenies for testing. Testing usually is not conducted in winter nurseries because environments are quite different. Table 7.34 illustrates some possible combinations of methods for use of the two seasons for different selection methods.

For mass and half-sib I selection, it is not possible to use winter nurseries because selection and recombination is completed in each summer season. For other methods of selection, winter seasons can be used to reduce the number of years to complete each cycle of selection. Use of winter seasons is dictated by availability of data for making selections, sequence of selection programs, and funds available. If the harvest is delayed for any reason, the winter planting may be too late to have seed available for planting in the next summer season. Winter seasons are useful to reduce cycle intervals, but researchers must be flexible in their use because unreasonable conditions may indicate making adjustments in planned selection programs. For example, full-sib progenies also can be produced at time of recombination to complete one cycle in 2 years.

Variance component estimates used to calculate expected gain for different selection methods are shown in Table 7.35. Three traits that have different heritability estimates were chosen: 38.9, 68.4, and 92.0% on a progeny mean basis for yield, ear length, and ear height, respectively, and 8.4, 19.9, and 59.8% on an individual plant basis. Ear height has a much greater heritability than yield, and mass selection would be expected to be much more effective for ear height than for yield. It seems that some type of progeny evaluation would be considerably better than mass selection for yield improvement, but we must consider the number of years required to complete each cycle of selection. Hence gain per cycle may be greater by some type of progeny evaluation but total gain may be greater by mass selection because one cycle of selection can be completed each summer session.

Expected gain for different selection methods for the three traits of Iowa Stiff Stalk Synthetic are given in Table 7.36. Expected gain is presented on both per-year and per-cycle basis to illustrate the effect of years on predicted gain for one cycle. In most instances the greater the number of years per cycle the greater the expected gain. Expected gain per year, however, is reduced when the cycle interval is increased. Effectiveness of different selection methods is illustrated by comparing expected gain for yield and ear height. Expected gain for yield per year by mass selection is considerably less than for other selection methods. Heritability estimates for yield show that the estimate on progeny mean basis is 4.6 times greater
<table>
<thead>
<tr>
<th>Season</th>
<th>Mass</th>
<th>Half-sib I&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Half-sib II&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Half-sib III&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Full-sib</th>
<th>S&lt;sub&gt;1&lt;/sub&gt;</th>
<th>S&lt;sub&gt;2&lt;/sub&gt;</th>
<th>Inbred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Produce S&lt;sub&gt;1&lt;/sub&gt; progenies</td>
<td>Produce S&lt;sub&gt;1&lt;/sub&gt; progenies</td>
<td>Produce S&lt;sub&gt;1&lt;/sub&gt; progenies</td>
</tr>
<tr>
<td>Summer</td>
<td>Selection</td>
<td>Selection</td>
<td>Produce HS progenies</td>
<td>Produce HS progenies, self</td>
<td>—</td>
<td>Test and selection</td>
<td>Produce S&lt;sub&gt;2&lt;/sub&gt; progenies</td>
<td>—</td>
</tr>
<tr>
<td>Winter</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>—</td>
</tr>
<tr>
<td>Summer</td>
<td>Selection</td>
<td>Selection</td>
<td>Test and selection</td>
<td>Test and selection</td>
<td>Recombine</td>
<td>Test and selection</td>
<td>Produce S&lt;sub&gt;3&lt;/sub&gt; progenies</td>
<td>—</td>
</tr>
<tr>
<td>Winter</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>Recombine</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>—</td>
</tr>
<tr>
<td>Summer</td>
<td>Selection</td>
<td>Selection</td>
<td>Produce HS progenies</td>
<td>Produce HS progenies, self</td>
<td>—</td>
<td>Test and selection</td>
<td>Recombine</td>
<td>—</td>
</tr>
<tr>
<td>Winter</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>—</td>
</tr>
<tr>
<td>Summer</td>
<td>Selection</td>
<td>Selection</td>
<td>Test and selection</td>
<td>Test and selection</td>
<td>Recombine</td>
<td>—</td>
<td>Recombine</td>
<td>—</td>
</tr>
<tr>
<td>Winter</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>Recombine</td>
<td>—</td>
<td>—</td>
<td>Recombine</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>a</sup>Half-sib I is modified ear-to-row selection; half-sib II is half-sib selection with recombination of remnant half-sib seed; and half-sib III is half-sib selection with recombination of selfed seed
Table 7.35 Component of variance estimates for three traits of Iowa Stiff Stalk Synthetic obtained from 800 full-sib progenies evaluated across six environments (Silva, 1974)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Estimates of components of variance</th>
<th>$\hat{h}^2$ (%)</th>
<th>Mean$^a$</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\sigma}_w^2$</td>
<td>$\hat{\sigma}^2$</td>
<td>$\hat{\sigma}_{DE}^2$</td>
<td>$\hat{\sigma}_{AE}^2$</td>
</tr>
<tr>
<td>Yield (g/plant)</td>
<td>1301 ± 18</td>
<td>185 ± 7</td>
<td>75 ± 12</td>
<td>92 ± 10</td>
</tr>
<tr>
<td>Ear length (cm)</td>
<td>3.98 ± 0.06</td>
<td>0.64 ± 0.02</td>
<td>0.26 ± 0.04</td>
<td>0.22 ± 0.03</td>
</tr>
<tr>
<td>Ear height (cm)</td>
<td>77.9 ± 1.3</td>
<td>14.0 ± 0.6</td>
<td>7.8 ± 0.9</td>
<td>8.0 ± 0.9</td>
</tr>
</tbody>
</table>

$^a$Heritability estimates calculated on a progeny mean basis as $\hat{h}^2 = \hat{\sigma}_A^2 / (\hat{\sigma}_w^2/8 + \hat{\sigma}_{DE}^2/4 + \hat{\sigma}_{AE}^2/4 + \hat{\sigma}_D^2 + \hat{\sigma}_A^2) \times 100$
### Table 7.36: Expected gain per year and per cycle for eight methods of recurrent selection in Iowa Stiff Stalk Synthetic for three traits

<table>
<thead>
<tr>
<th>Method of selection</th>
<th>Coefficient$^b$</th>
<th>Seasons per cycle</th>
<th>Expected gain$^a$</th>
<th>Yield (q/ha)</th>
<th>Ear length (cm)</th>
<th>Ear height (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
<td>0.54</td>
<td>0.54</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>Mass</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>17.84</td>
<td>17.84</td>
</tr>
<tr>
<td>Half-sib I$^c$</td>
<td>$\frac{1}{8}$</td>
<td>1</td>
<td>1.75</td>
<td>1.75</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>Half-sib II$^c$</td>
<td>$\frac{1}{4}$</td>
<td>2</td>
<td>1.75</td>
<td>3.50</td>
<td>0.46</td>
<td>0.92</td>
</tr>
<tr>
<td>Half-sib III$^c$</td>
<td>$\frac{1}{2}$</td>
<td>3</td>
<td>2.34</td>
<td>7.01</td>
<td>0.61</td>
<td>1.83</td>
</tr>
<tr>
<td>Full-sib</td>
<td>$\frac{1}{2}$</td>
<td>2</td>
<td>2.25</td>
<td>4.51</td>
<td>0.63</td>
<td>1.26</td>
</tr>
<tr>
<td>S$_1$ progeny</td>
<td>1</td>
<td>2</td>
<td>3.61</td>
<td>7.22</td>
<td>0.96</td>
<td>1.92</td>
</tr>
<tr>
<td>S$_1$ progeny$^d$</td>
<td>1</td>
<td>3</td>
<td>2.41</td>
<td>7.22</td>
<td>0.64</td>
<td>1.92</td>
</tr>
<tr>
<td>S$_2$ progeny</td>
<td>$\frac{3}{2}$</td>
<td>3</td>
<td>3.28</td>
<td>9.85</td>
<td>0.82</td>
<td>2.46</td>
</tr>
<tr>
<td>S$_2$ progeny$^d$</td>
<td>$\frac{3}{2}$</td>
<td>4</td>
<td>2.46</td>
<td>9.85</td>
<td>0.62</td>
<td>2.46</td>
</tr>
<tr>
<td>Inbred line</td>
<td>2</td>
<td>5</td>
<td>2.34</td>
<td>11.73</td>
<td>0.56</td>
<td>2.81</td>
</tr>
</tbody>
</table>

$^a$ Expected gain calculated assuming four environments, each including two replications

$^b$ See Table 6.12

$^c$ See Table 7.34

$^d$ Additional season of recombination is used before developing the S$_1$ progenies
than for individual plants. Heritability estimates for ear height are greater and the progeny mean estimate (92.0%) is 1.5 times greater than the individual plant estimate (59.5%). Consequently, expected gain of mass selection for ear height per year is greater than all other selection methods except S₁ progeny (2 years). If selection for ear height can be made before flowering, undesirable male plants can be detasseled and expected gain from mass selection is increased (the coefficient is increased from 0.5 to 1) and is greater than all other selection methods on a per-year basis. (Note: Usually selection is for lower rather than higher ear heights in maize breeding.) Expected gain from mass selection for yield improvement and increased ear length also could be increased by controlling the male parents, but present techniques do not provide for a direct identification of the superior yielding and longer eared plants at flowering time. Expected gains listed in Table 7.35 apply only when testing is conducted in four environments with two replications in each environment.

Expected gains for different selection methods in Table 7.36 are for a specific situation such as the US Corn Belt. In tropical and semitropical environments, the frost-free period may extend throughout the year. Consequently the seasons may not differ very much during the year, or they may be quite different because of some specific weather factor. It may be possible to get three crop seasons per year as far as temperatures and light durations are concerned; but rainfall may be sparse in one season, thus making yield trials impossible. In that case it may be possible to irrigate a small area for making recombination of recurrent selection programs. Hence, efficiency of selection in tropical and semitropical environments may be enhanced by effective use of different seasons available. Early-maturing germplasm (<90RM) can also allow two winter generations and a summer one in the northern USA for a total of three generations per year as used, in some instances, by the NDSU corn breeding program.

S. A. Eberhart (personal communication, 1971) calculated the expected gain in yield for different selection methods for situations having different numbers and types of seasons. Estimates of parameters used to calculate expected gain were obtained from four half-sib selection trials in Kenya, East Africa (Table 7.37). Three situations (A, B, and C) were used to predict gain for five different combinations of seasons that represent different maize growing areas. Additive genetic variance was constant for situations A, B, and C but the other parameters were varied to show their effects on predicted gain. The relative changes in predicted gains for situations A, B, and C (Table 7.36) apply only when yield trials are conducted with two replications at four environments within the same season. The relative changes in predicted gains will not be the same for all methods if numbers of replications and locations are changed. When the seasons are similar, however, relative gains among selection methods will not change whether there are one, two, or three seasons per year.

Gain per cycle for different selection methods (Table 7.38) depends on heritability of the trait and type of progeny evaluation used. Mass selection has the smallest expected gain per cycle for situations A and B but mass selection only requires one growing season to complete a cycle. With a greater heritability value (50%) in situation C, expected gain per cycle of mass selection exceeds half-sib I and
### Table 7.37 Estimates of genetic parameters for yield for comparing selection methods

<table>
<thead>
<tr>
<th>Situation</th>
<th>$\hat{\sigma}^2_A$</th>
<th>$\hat{\sigma}^2_D$</th>
<th>$\hat{\sigma}^2_{AL}$</th>
<th>$\hat{\sigma}^2_{DL}$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{\sigma}^2_{me}$</th>
<th>$\hat{\sigma}^2_w$</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>60</td>
<td>30</td>
<td>68</td>
<td>34</td>
<td>98</td>
<td>967</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>60</td>
<td>0</td>
<td>68</td>
<td>0</td>
<td>98</td>
<td>967</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>60</td>
<td>50.0</td>
</tr>
</tbody>
</table>

a $\hat{\sigma}^2$, $\hat{\sigma}_{me}^2$, and $\hat{\sigma}_w^2$ refer to the experimental error, micro-environmental interaction, and within-plot variance, respectively.

b Estimates were obtained from 4 half-sib experiments with 21 plants per plot, 2 replications, and 4 environments from analysis pooled over 3 years. Estimates of dominance variance and interactions were assumed to be half of their corresponding additive genetic variance parameters. The B and C situations are arbitrary to show the effect of changing parameters with additive genetic variance held constant.

II selection. Genetic gain per year depends also on types of seasons available each year. For low heritability (situations A and B) and similar seasons, full-sib, S1, and S2 selection give the greatest predicted gain per year. If one has one or two similar seasons per year and high heritability (situation C), mass selection is clearly superior to other methods. With no dominance variance, low heritability (situation B), and similar seasons, full-sib selection gives as much gain as S1 and S2 progeny selection but yield trials are required every second season for full-sib selection. Half-sib I has the greatest predicted gain with two seasons, particularly when heritability is high. With two non-similar seasons per year, full-sib selection has greater predicted gain per year than other selection methods, particularly for situations B and C that have no dominance variance. Two non-similar seasons represent conditions used in calculating predicted gain in Table 7.36 for Iowa Stiff Stalk Synthetic. However, the predicted gain for full-sib selection was less than for S1 and S2 progeny selection because dominance variance was slightly greater than additive genetic variance for yield. The effect of dominance variance on predicted gain for full-sib selection relative to S1 and S2 progeny selection for two non-similar seasons is also shown in Table 7.38 for situation A.

With three non-similar seasons in 2 years, S1 progeny selection is predicted to give more gain than S2 progeny selection. Predicted gain for three non-similar seasons per year is greater for half-sib III and S1 progeny selection than for other selection methods; half-sib III and S1 progeny selection are about the same for situation C but S1 progeny selection has greater predicted gain for situations A and B, which have low heritability values.

Choice of a selection method depends on the trait under selection, type of population in which selection is initiated, and objectives of the selection program in relation to the total breeding program. If an exotic population is introduced into the breeding program, mass selection probably will be effective in adapting the population to local conditions (Hallauer and Sears, 1972; Hallauer and Carena, 2009). As shown in Table 7.36, highly heritable traits (ear height, in this instance)
Table 7.38  Predicted gain per cycle and per year for three sets of genetic parameters (A, B, and C)\textsuperscript{a} and for five combinations of growing seasons

<table>
<thead>
<tr>
<th>Selection method</th>
<th>Seasons/cycle</th>
<th>Gain/cycle</th>
<th>Two similar seasons</th>
<th>Two non-similar seasons</th>
<th>Three non-similar seasons in 2 years</th>
<th>Three non-similar seasons/year</th>
<th>One season/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>1</td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
<td>3.0</td>
<td>3.2</td>
<td>13.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
<td>1.5</td>
<td>1.6</td>
<td>6.8</td>
</tr>
<tr>
<td>Half-sib I\textsuperscript{b}</td>
<td>1</td>
<td>2.3</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
<td>4.6</td>
<td>6.8</td>
</tr>
<tr>
<td>Half-sib II\textsuperscript{b}</td>
<td>2</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
<td>4.7</td>
<td>4.7</td>
<td>6.7</td>
</tr>
<tr>
<td>Half-sib III\textsuperscript{b}</td>
<td>3</td>
<td>9.4</td>
<td>9.4</td>
<td>13.4</td>
<td>6.2</td>
<td>6.2</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.4</td>
<td>9.4</td>
<td>13.4</td>
<td>6.2</td>
<td>6.2</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.4</td>
<td>9.4</td>
<td>13.4</td>
<td>6.2</td>
<td>6.2</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.4</td>
<td>9.4</td>
<td>13.4</td>
<td>6.2</td>
<td>6.2</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.4</td>
<td>9.4</td>
<td>13.4</td>
<td>6.2</td>
<td>6.2</td>
<td>8.9</td>
</tr>
<tr>
<td>Full-sib</td>
<td>2</td>
<td>6.8</td>
<td>7.3</td>
<td>9.5</td>
<td>6.8</td>
<td>7.4</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.8</td>
<td>7.3</td>
<td>9.5</td>
<td>6.8</td>
<td>7.4</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.8</td>
<td>7.3</td>
<td>9.5</td>
<td>6.8</td>
<td>7.4</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.8</td>
<td>7.3</td>
<td>9.5</td>
<td>6.8</td>
<td>7.4</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.8</td>
<td>7.3</td>
<td>9.5</td>
<td>6.8</td>
<td>7.4</td>
<td>9.5</td>
</tr>
<tr>
<td>S\textsubscript{1}</td>
<td>3</td>
<td>10.6</td>
<td>11.1</td>
<td>13.5</td>
<td>7.0</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.6</td>
<td>11.1</td>
<td>13.5</td>
<td>7.0</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.6</td>
<td>11.1</td>
<td>13.5</td>
<td>7.0</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.6</td>
<td>11.1</td>
<td>13.5</td>
<td>7.0</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.6</td>
<td>11.1</td>
<td>13.5</td>
<td>7.0</td>
<td>7.4</td>
<td>9.0</td>
</tr>
<tr>
<td>S\textsubscript{2}</td>
<td>4</td>
<td>13.6</td>
<td>13.9</td>
<td>16.6</td>
<td>6.8</td>
<td>7.0</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.6</td>
<td>13.9</td>
<td>16.6</td>
<td>6.8</td>
<td>7.0</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.6</td>
<td>13.9</td>
<td>16.6</td>
<td>6.8</td>
<td>7.0</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.6</td>
<td>13.9</td>
<td>16.6</td>
<td>6.8</td>
<td>7.0</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.6</td>
<td>13.9</td>
<td>16.6</td>
<td>6.8</td>
<td>7.0</td>
<td>8.3</td>
</tr>
</tbody>
</table>

\textsuperscript{a}A, B, and C as defined in Table 7.36

\textsuperscript{b}See Table 7.35
can be effectively changed by mass selection, which also is important for adapting day-length-sensitive germplasm to the US Corn Belt. For traits such as days to flower, prolificacy, and ear height, effectiveness of mass selection can be enhanced by detasseling the male plants before anther dehiscence because parental control is increased. However, partial parental control can be expected by pollen shed from earlier maturing plants.

Half-sib I is a modification that seems promising: (1) the method is simple and progeny data are available for yield evaluation and (2) one cycle can be completed each year (or season). If the selection program is an adjunct to an applied breeding program, half-sib III, S1, and S2 progeny selection may be favored because preliminary yield test data can be obtained from the evaluation trials. For half-sib III, testcross information is available and S1 lines used for recombination can be included in the applied breeding program for additional inbreeding and testing.

If genetic variation is primarily additive S1 and S2 selection should be effective, and both generate new lines for the applied breeding program. The S2 selection requires an additional season to complete a cycle and predicted gain per year may be less than for S1 selection. The slightly smaller predicted gain, however, may be offset by selection advantages for an applied breeding program. For instance, if 500 S1 progenies are saved from selected S0 plants of some population, the S1 progenies can be screened for pest resistance, seedling vigor, emergence percentage, lodging, seed set, and other phenotypic plant traits in a two-replication test in one environment. Selected S1 plants within and among the S1 progenies can be advanced to the S2 generation in a breeding and/or disease nursery. Further selection at harvest may reduce the number of S2 progenies to 100 for testing. If 20–25 S2 progenies are selected for recombination, intense selection pressure has been given for highly heritable traits before expensive yield tests are conducted. This illustration is one example of how the selection method can be adapted for an applied breeding program that has the primary objective of developing new inbred lines. Innovations by breeders can be made in all selection methods to fit their breeding programs and objectives.

Predicted gains given are for only one cycle. If selection is effective in changing allele frequencies of the original population, estimates of genetic parameters of the original populations may be invalid for use in predicting gain in future cycles. These markers need to be identified in each selection cycle. For progeny evaluation trials, however, estimates of genetic parameters can be obtained in each cycle and can be used to predict gain expected in the next cycle. Pooling estimates from successive cycles of selection will provide better estimates of parameters used in predicting gain provided the change in allele frequency is not great. If significant changes in allele frequency are suspected, estimates from the most recent cycle of selection should be used even though they may have greater errors than pooled estimates from several cycles.

Marker-assisted selection (MAS) with the purpose to enhance elite germplasm and develop new cultivars (e.g., inbred lines and hybrids) is still in its infancy. Several theories and hypotheses (e.g., simulation studies) proposing MAS for genomic segments affecting quantitative traits remain to be tested. Few research
studies have compared MAS with phenotypic selection based on progeny testing and relative comparisons were often made via computation simulation with an additive genetic model. Very few attempts to complement mating design and molecular studies have been reported. Complementary approaches could be used for specific traits that are still difficult to measure (e.g., root traits, rate of dry down) and an optimal balance with MAS should be targeted (Hallauer and Carena, 2009) for greater short- and long-term responses. Maize breeders emphasize breeding and selection strategies to improve current materials, which have been developed by previous long-term selection to enhance genetic complexes. Use of half-sib and full-sib family RRS methods would require molecular markers for two populations and every selection cycle, and how they would combine with the opposing population. Therefore, for intra- and inter-population recurrent selection methods molecular markers would need to be identified and validated for each cycle of selection obtained (Hammond and Carena, 2008). The maize genome sequence is nearly completed and selection efficiency of phenotypes could be complemented with gene information to enhance selection for quantitative traits. However, the current sequence is based on B73 and knowledge on unique and rare alleles is desirable as each genotype has its own genetic effects.

Recurrent selection has demonstrated to be effective across methods. To maximize genetic improvement and contribute to future significant genetic advances, selection should be conducted on a continuous basis. In genetically broad-based maize populations, different schemes of recurrent selection can be used to improve grain yield. Additive genetic effects usually are of greatest importance but non-additive effects (dominance, overdominance, and epistasis) become important to enhance the heterotic groups for inbred line and hybrid development. Prediction of both types of genetic effects and continuous deployment of useful genetic diversity would be desirable for continued incremental genetic gains in the future.

References


References


References


Selection: Experimental Results


Paterniani, E. 1968. Avaliação do método de seleção entre e dentro de famílias de meios irmãos no melhoramento do milho (*Zea mays L.*). Tese de Cátedra, ESALQ-USP, Piracicaba, Brazil.


The inbred–hybrid concept was created in the public sector with the direct influence of Darwin, Festetics, Mendel, and Vilmorin. East related those biological principles to the more practical plant improvement studies to achieve his goals (Hayes, 1956). The progeny test was defined by Allard (1960) as ‘a test of the value of a genotype based on the performance of its offspring produced in some definite system of mating.’ It was used as early as 1850 by Vilmorin in France, and it proved to be a highly effective procedure for the improvement of sugar content of sugar beets ($Beta vulgaris$). This method of line selection with progeny testing was known as the ‘Vilmorin method’ or ‘Vilmorin isolation principle’ and was introduced in several plant breeding programs in the latter part of the 19th century. The progeny test in maize was first used in 1896 by Hopkins, starting the well-known program for half-sib recurrent selection of maize oil and protein content (e.g., the ear-to-row half-sib selection procedure).

Davis (1927) suggested the use of the testcross procedure, which is a type of progeny test, to evaluate the combining ability of inbred lines in a hybrid maize breeding program. After Jenkins and Brunson (1932) reported on the effective use of the testcross procedure, it was widely adopted in breeding programs. Other types of progeny tests, such as the ones based upon full-sib and/or inbred ($S_1$ or $S_2$) progenies, were subsequently introduced.

After Sprague and Tatum (1942) introduced the concepts of general combining ability (GCA) and specific combining ability (SCA), new approaches for the use of progeny test or testcross were suggested. Thus recurrent selection methods (Jenkins, 1940; Hull, 1945; Comstock et al., 1949; Lonnquist, 1949) were introduced, widening the horizon for population improvement in maize. In selection for GCA, it was proposed to utilize a broad base heterogeneous population as tester. It can be either the parental population or any broad genetic base (synthetic or open-pollinated variety), unrelated population. In all instances genotypes are tested with a representative sample of genotypes in the tester, that is, each plant in the base population is crossed with a random sample of gametes from the tester. Each testcross, therefore, is a type of half-sib family. When the tester has a narrow genetic base (inbred line or single cross), selection among testcrosses was proposed to be for SCA. The original proposal of reciprocal recurrent selection (RRS) is similar to selection for GCA because
two populations are selected simultaneously, with the testcrosses for each one using the opposite population as the tester.

The use of testcross (or topecross) in maize breeding has one of the following objectives:

1. evaluation of combining ability of inbred lines in a hybrid breeding program or
2. evaluation of breeding values of genotypes (plants) for population improvement.

In each instance a problem of choice of tester is essentially the same. The goal is to find a tester that provides the best discrimination among genotypes according to the purposes of selection.

For inbred line evaluation, a desirable tester was defined by Matzinger (1953) as one that combines the greatest simplicity in use with the maximum information on performance to be expected from tested lines when used in other combinations or grown in other environments. It was recognized, however, that no single tester can completely fulfill these requirements. Rawlings and Thompson (1962) defined a good tester as one that classifies correctly relative performance of lines and discriminates efficiently among lines under test. For improvement of breeding populations, Allison and Curnow (1966) defined the best tester as one that maximizes the expected mean yield of the population produced from random mating of selected genotypes. Hallauer (1975) pointed out that in general a suitable tester should include simplicity in use, provide information that correctly classifies the relative merit of lines, and maximize genetic gain.

After Jones (1918) suggested the double-cross hybrid procedure, selection among hybrids was based on the cross performance of a set of inbred lines. With relatively few inbred lines developed in hybrid programs in the period 1920–1930, the \( n(n - 1)/2 \) combinations possible from a set of \( n \) inbred lines could be tested for the evaluation of their cross performance. As the number of lines increased, crossing and testing inbred lines in all combinations became impossible. Evaluation of inbred lines themselves had little value because of the inconsistency of correlation between characters of the inbred lines and their performance in \( F_1 \) hybrid crosses (e.g., poor lines could give excellent hybrids and vice versa). The testcross test introduced by Davis (1927) made possible the screening of inbred lines based first on GCA with a broad base tester. This procedure was shown to be effective by Jenkins and Brunson (1932) and was widely used subsequently. Johnson and Hayes (1936) also reported that inbred lines giving high yields in testcrosses were more likely to produce better single crosses. This is in agreement with later studies on correlations between early- and late-generation testcrosses (Jensen et al., 1983; Lile and Hallauer, 1994) and predictions based on molecular markers (Eathington et al., 1997; Johnson, 2004).

Early testing of inbred lines was suggested by Jenkins (1935) and Sprague (1939, 1946). Early testing differed in two main respects from the usual procedure for testing inbred lines:
(1) $S_0$ plants were crossed with a tester at the time of first selfing and combining ability and general performance of the testcross progeny were determined.

(2) The first discarding allowed a greater concentration of efforts on the families of greatest promise during the $S_1$ and $S_2$ generations where greater opportunities for selection within lines existed (Sprague, 1946).

Several subsequent reports showed that the combining ability of lines had a reasonably good stability through the generations of inbreeding, as indicated in reviews presented by Loeffel (1964, 1971). In general, development of inbred lines is a process of sequential selection because some lines are discarded early by their poor performance both in early testing and across nurseries and by later results for general and for specific combining ability. However, the recent increase in production of inbred lines through doubled haploids addresses some of the stability concerns but generates more unknowns (Hallauer and Carena, 2009). If early testing is one of the choices for an applied breeding program, Jenkins (1935) stated that the combining ability of a line is established early in the inbreeding process and remained relatively stable in succeeding generations of inbreeding. In other cases, however, the genetic variability among and within lines during successive generations of inbreeding may cause early testing to eliminate lines that may have different and better combining abilities as homozygous lines. With early testing the effects of selection during inbreeding might be reduced and the large number of experimental field plots that would be required to evaluate the early inbred-generation lines might approach those needed for doubled haploids. Selection during inbreeding process could be continued but only among those lines that had above average combining ability. Breeders can continue to practice visual selection among and within lines for desired plant types, seed set, maturity, etc., that have relatively higher heritability estimates. Also, the breeder can use certain populations for early testing and/or doubled-haploid inbred line development while other types of populations can be subjected to other methods depending on their genetic structure.

In either early or late testing, the most questionable point is the choice of a tester to evaluate the combining ability. In some instances when the objective is the replacement of a line in a specific combination, SCA is of prime importance and the most appropriate tester is the opposite inbred line parent of a single cross or the opposite single-cross parent of the double cross (Matzinger, 1953). Often the choice of the best tester is still obscure.

Green (1948a) compared the relative value of two testers, a low-yielding, lodging-susceptible variety and a high-yielding, lodging-resistant, double-cross hybrid. Because the two testers ranked individual segregates of $F_2$ plants differently and because there was a difference between testers in average testcross yields, the existence of striking differences in their genetic structures was suggested. It was not possible to determine from available data which tester gave the better estimate of average combining ability; it was suggested that the best estimate would be obtained using the average performance of testcrosses with both testers rather than testcrosses of either tester alone. It was also suggested that a synthetic variety composed of the lines in current use could be employed to measure combining ability of new inbred
material. Keller (1949) also found a very low correlation for agronomic characters of related tester (parental population) crosses and the same characters in the unrelated tester (single-cross R4 × Hy), indicating that testers did not rank the lines in similar order. No evidence was presented for the superiority of one over the other.

Recurrent selection experiments have shown successful results for improving populations in favor of early-generation testing. These experiments also indicated differences between testers in ranking the genotypes of a population (Lonnquist and Rumbaugh, 1958; Horner, 1963; Lonnquist and Lindsey, 1964; Horner et al., 1969). In selecting for GCA, the only apparent requirement was that the tester parent be genetically heterogeneous (with a broad genetic base). The GCA of a line, however, is not a fixed property of a line but depends on the genetic composition of the population with which it has been crossed (Kempthorne and Curnow, 1961).

### 8.1 Theory

The relative value of testers was theoretically examined by Hull (1945). His conclusions were based on the variances among and within testcrosses considering one locus with two alleles distributed as 1AA:2Aa:1aa and with complete dominance and arbitrary values for genotypic effects (Table 8.1).

If inbred lines are drawn from a population, their frequencies will be \( \frac{1}{2} AA \) and \( \frac{1}{2} aa \). When they are testcrossed with the parental population the mean of testcrosses will be 1 and \( \frac{1}{2} \), respectively. Thus the general mean is \( \frac{3}{4} \). The variance among testcrosses is \( \frac{1}{16} \) and consequently the variance within testcrosses is \( \frac{3}{16} - \frac{1}{16} = \frac{2}{16} \). Hull (1945, 1946, 1952) thus stated that theoretically the most efficient tester would be one that is homozygous recessive at all loci and that homozygosity for dominance alleles at any locus should be avoided. Support for this conclusion was also made from the method of constant-parent regression, i.e., regression of performance of offspring on the performance of variable parents for a particular constant parent. Such a regression was shown to be largest with gene frequency 0 for the

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequency</th>
<th>Genotypic values</th>
<th>AA tester</th>
<th>aa tester</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>( \frac{1}{4} )</td>
<td>1</td>
<td>( \frac{1}{4} )</td>
<td>1</td>
</tr>
<tr>
<td>Aa</td>
<td>( \frac{1}{4} )</td>
<td>1</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>Aa</td>
<td>( \frac{1}{4} )</td>
<td>1</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>aa</td>
<td>( \frac{1}{4} )</td>
<td>0</td>
<td>( \frac{1}{4} )</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>( \frac{3}{4} )</td>
<td>1</td>
<td>( \frac{1}{2} )</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>( \frac{3}{16} )</td>
<td>0</td>
<td>( \frac{2}{16} )</td>
<td></td>
</tr>
</tbody>
</table>
character in the constant parent. Regression would be zero either for gene frequency 1 for complete dominance or for gene frequency at equilibrium for overdominant gene action. Some empirical results showed that the regression of $F_1$ on parental lines was either essentially zero or negative with strong testers and that weaker lines of several samples as testers gave a regression of the order of 0.6 and higher. Results reported by Green (1948a) were in accordance with Hull’s hypothesis because the lodging-susceptible tester showed a much greater range for standability than the lodging-resistant tester. On the other hand, Keller (1949) found no difference in the line × tester interaction variance using two groups (high yielding and low yielding) of testers. Under Hull’s hypothesis it was expected that the low-yielding group would give a higher interaction variance. The lack of interaction would favor the common practice of breeders of using unrelated high-yielding testers. In contrast to earlier years, when unique heterotic groups were not defined, the choice of testers is usually elite inbred lines from the opposite heterotic group.

### 8.1.1 Variance Among Testcrosses

Rawlings and Thompson (1962) added a significant contribution to the study of testers. They used 6 inbred lines, grouped according to their GCA (low, intermediate, and high), crossed with 10 heterozygous strains (five paired crosses of high-yielding selections HH and five paired crosses of low-yielding selections LL) obtained from a divergent recurrent selection program for yield in synthetic A. Comparisons were made considering the heterozygous strains as testers for the six inbred lines (case I) and, conversely, considering the inbred lines as testers for the heterozygous strains (case II). In case I, with few exceptions, ranking of the lines corresponded to that expected on the basis of the assumed levels of GCA for yield, for both high-yielding and low-yielding heterozygous strains as testers. In case II, the correct relative classification was given on average by each of the inbred lines as testers. Considerable overlapping of the ranges of the two groups occurred, but in both instances the testers seemed to give a reasonably correct ranking of the genotypes. The authors also pointed out that a second requirement in a good tester is precision in discriminating among genotypes under test; that is, the best tester would be the one that would give most precise classification among entries for a given amount of testing, thus allowing the testing of a greater number of entries for a given degree of precision of estimates. They used the $F$-test for among-entries variance $\hat{\sigma}_E^2$ to compare relative efficiency of testers for discriminating among genotypes. In case I, the $F$-test for the hypothesis $\hat{\sigma}_E^2 = 0$ showed that four testers (one of HH group and three of LL group) gave significant differences among inbred lines. In case II, inbred lines used as testers (two high, one intermediate, and one low in GCA) measured the differences among the 10 intercrosses as significant. Using the Shuman and Bradley (1957) test of significance for relative sensitiveness of two methods of measurement, they found that only one of the possible comparisons for sensitivity (that involving the inbred lines NC44 and NC216) was significant. On the other hand, average sensitivity for the LL group in case I was 28% greater.
than average sensitivity of the HH group of heterozygous strains used as testers. In case II, average sensitivity of intermediate and low-combining testers was 98 and 90%, respectively, greater than average sensitivity of high-combining testers. Overall results were in favor of the theory that low-performing testers, presumably with low frequency of favorable alleles at important loci, are the most effective.

Theory supporting Rawlings and Thompson’s (1962) results was basically an extension of Hull’s hypothesis. They assumed a model with no epistasis and examined genetic variability among testers for different levels of dominance and different gene frequencies in the testers. After adaptation to the notation used in Chapter 2, genetic variance among testcross progenies for one locus and for a particular tester can be expressed as

$$\hat{\sigma}_t^2 = \left(\frac{1}{2}\right) p (1 - p) (1 + F) [a + (1 - 2r) d]^2$$

which gives total genetic variance after summation over all loci. For a given population with fixed values of $p$, $F$, $a$, and $d$, it is clear that variance among testcrosses will depend on gene frequency $r$ in the tester. With no dominance ($d = 0$), genetic variance among testcrosses is constant for any gene frequency in the tester:

$$\hat{\sigma}_t^2 = \left(\frac{1}{2}\right) p (1 - p) (1 + F) a^2$$

For any level of positive dominance, genetic variance for tester gene frequency $r = 0$ will always be greater. For gene frequency $r = 0.5$ in the tester, $\hat{\sigma}_t^2$ has a constant value for any level of dominance. For gene frequency $r = 1.0$ in the tester, $\hat{\sigma}_t^2$ is zero for complete dominance and equals the value for $r = 0.5$ for levels of dominance either 0 or 2. Such relations are presented graphically in Fig. 8.1 from results of Rawlings and Thompson (1962).

Figure 8.1 shows that a low gene frequency in the tester will give greater variances in the range of partial to complete dominance of genes. High gene frequency in the tester may give greater variances if overdominance is of considerable importance. Several reports have shown evidence that genes controlling yield and other important traits in maize are mostly additive and dominance ranges from partial to complete (see Chapter 5). Therefore, the hypothesis under which a low-yielding tester is better than a high-yielding tester is more likely to be true. Although Rawlings and Thompson’s results showed a general trend in favor of their hypothesis, they were considered not conclusive by the authors because of lack of significance in most comparisons relative to efficiency of detecting differences among testers. It was recognized that epistasis might completely confound results expected on the basis of only an additive dominance model.

The choice of tester is currently related to heterotic groups and ultimately to the hybrid product (Hallauer and Carena, 2009). Hence, the initial and advanced testing of new inbred lines will always include testers that are inbred lines that represent elite germplasm of the breeding program. Priority should be given to testers from the opposite group over poor testers for the same amount of genetic variability (e.g., elite industry testers). Choice of tester also depends on breeding goals. If a breeding
Fig. 8.1 Relative genetic variance among testcrosses for three tester gene frequencies and different levels of dominance

Goal is to improve the female parent of a successful hybrid (Dudley, 1982), then the male parent of the hybrid would be the logical choice of tester. If a new hybrid is the goal then inbred lines may be tested across breeding stages with different elite inbred line testers representing opposite heterotic groups. Barata and Carena (2006) suggested the possibility of developing unique inbred lines with good combining ability across testers.

8.1.2 Expected Changes in Gene Frequency

Comstock (1964) used the same model to analyze the tester problem by comparing expected changes in gene frequency ($\Delta p$) after selection among testcrosses for different testers. It was shown that

1) $\Delta p$ is proportional to $a + (1 - 2p)d$ (according to the notation in Chapter 2) when the parental population is used as tester and

2) $\Delta p$ is proportional to $a + (1 - 2r)d = a + (1 - 2p)d + 2(p - r)d$ when an unrelated population is used as tester.

Thus it is clear that change in gene frequency $\Delta p$ is expected to be greater using an unrelated tester whenever the quantity $p - r$ is on the average positive and of sufficient magnitude than when using the parental population as the tester. This situation would be more likely to occur with a poor tester. Therefore, with low gene
frequency at important loci a poor performance tester is used rather than a good performance tester.

Lopez-Perez (1979) conducted a comprehensive study that included 50 unselected S₁ and 50 unselected S₈ lines crossed with five testers that were selected for their expected differences in gene frequencies for yield. The source population for the unselected sample of lines and four of the five testers was Iowa Stiff Stalk Synthetic (BSSS). The samples of 50 lines were from a larger group of 250 S₁ and 247 S₈ lines developed by single-seed descent from Iowa Stiff Stalk Synthetic (Hallauer and Sears, 1973). The 50 S₈ lines were direct descendants of the 50 S₁ lines. The five testers were BSSS, the parental source population; BS13(S)C1, an improved strain of BSSS (Hallauer and Smith, 1979); BSSS-222, an S₈ line derived from BSSS and identified as one of the poorest yielding lines per se; B73, a high-performance line derived from BS13(HT) after five cycles of half-sib recurrent selection; and Mo17, an unrelated line developed by pedigree selection from the cross 187-2 × C103. The 500 testcrosses (50 S₁ lines and 50 S₈ lines crossed with five testers) were evaluated in five experiments to estimate and compare the variability among the testcrosses for each tester. Estimates of components of variance of S₁ and S₈ lines and their interactions with the five testers are presented in Table 8.2 for six traits. Estimated components of variance among S₁ and S₈ lines agreed with expectations, assuming gene frequency was 0.5 or no dominance effects (see Chapter 2). The variation among S₈ lines was about two times the one among S₁ lines. Estimates of components of variance of S₁ lines and their interactions with testers were nearly equal, whereas interactions of S₈ lines with testers were less than the line component of variance. Hence, additive effects seemed more important than non-additive effects in causing differences among the S₁ and S₈ lines. Analyses of variance indicated that differences among lines and their interactions with testers were highly significant ($P \leq 0.01$) in all instances except for percentage of root lodging for S₁ lines, which was significant at the 5% level.

Analyses of variance showed that variation among testcrosses was highly significant for all testers except for B73 S₁ testcrosses. Table 8.3 shows estimates of components of variance for S₁ and S₈ testcrosses for each tester and, again, the evidence for larger variation among S₈ testcrosses.

### Table 8.2  Estimates of components of variance and standard errors for S₁ and S₈ lines and their interactions with testers

<table>
<thead>
<tr>
<th>Traits</th>
<th>Lines</th>
<th>Lines × testers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S₁</td>
<td>S₈</td>
</tr>
<tr>
<td>Yield</td>
<td>8.9 ± 2.5</td>
<td>23.3 ± 5.9</td>
</tr>
<tr>
<td>Grain moisture</td>
<td>0.5 ± 0.1</td>
<td>0.8 ± 0.2</td>
</tr>
<tr>
<td>Days to silk</td>
<td>0.7 ± 0.2</td>
<td>1.0 ± 0.3</td>
</tr>
<tr>
<td>Root lodging</td>
<td>1.2 ± 0.8</td>
<td>5.1 ± 1.7</td>
</tr>
<tr>
<td>Stalk lodging</td>
<td>11.0 ± 3.0</td>
<td>24.5 ± 6.7</td>
</tr>
<tr>
<td>Dropped ears</td>
<td>0.4 ± 0.1</td>
<td>0.9 ± 0.3</td>
</tr>
</tbody>
</table>
Table 8.3  Estimates of components of variance and standard errors for yield (q/ha) for S1 and S8 line testcrosses ($\hat{\sigma}^2_T$) and their interactions with environments ($\hat{\sigma}^2_{TE}$) for five testers

<table>
<thead>
<tr>
<th>Testers</th>
<th>$\hat{\sigma}^2_T$ (S1)</th>
<th>$\hat{\sigma}^2_T$ (S8)</th>
<th>$\hat{\sigma}^2_{TE}$ (S1)</th>
<th>$\hat{\sigma}^2_{TE}$ (S8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS</td>
<td>18.4 ± 6.5</td>
<td>41.7 ± 12.0</td>
<td>10.5 ± 6.2</td>
<td>21.4 ± 7.4</td>
</tr>
<tr>
<td>BS13(S)C1</td>
<td>10.6 ± 4.5</td>
<td>34.1 ± 10.2</td>
<td>3.7 ± 5.5</td>
<td>16.0 ± 6.8</td>
</tr>
<tr>
<td>BSSS-222</td>
<td>21.7 ± 6.4</td>
<td>38.7 ± 11.1</td>
<td>-6.9 ± 4.4</td>
<td>18.0 ± 7.1</td>
</tr>
<tr>
<td>B73</td>
<td>3.8 ± 3.5</td>
<td>26.4 ± 8.0</td>
<td>12.9 ± 6.5</td>
<td>6.2 ± 5.8</td>
</tr>
<tr>
<td>Mo17</td>
<td>25.9 ± 8.5</td>
<td>29.8 ± 9.5</td>
<td>17.2 ± 7.0</td>
<td>21.8 ± 7.5</td>
</tr>
<tr>
<td>Average</td>
<td>16.1 ± 5.9</td>
<td>34.1 ± 10.2</td>
<td>7.5 ± 5.9</td>
<td>16.7 ± 6.9</td>
</tr>
</tbody>
</table>

For the four related testers, estimates compare favorably with those expected relative to the choice of testers. BS13(S)C1 (the improved strain of BSSS) and B73 (a narrow-based elite line) had less variability among testcrosses than BSSS (the original population) and BSSS-222 (a poor performance line). Hence estimates indicate less variability among testcrosses for testers expected to have a greater frequency of favorable alleles within related testers. Mo17 was unrelated to the lines under test, and variability among S1 Mo17 testcrosses was similar to that for BSSS and BSSS-222, the two testers expected to have a lower frequency of favorable alleles. At the S8 level, estimates for Mo17 and B73 were similar but less than those for the other three testers. The average variation among S1 and S8 testcrosses had the same trend as the variation among S1 and S8 lines (Table 8.2), indicating that additive effects were primarily responsible for differences among testcrosses.

Estimates of the tester by environment interactions were, on the average, about one-half the variation among testcrosses (Table 8.3). At the S1 level, only testcrosses involving B73 and Mo17 had interactions with environments significantly different from zero, but the B73 testcrosses were the only group that had a non-significant interaction at the S8 level. Testcross by environment interactions at the S8 level averaged about twice those at the S1 level: 7.5 for S1 vs. 16.7 for S8. There were no trends among testers in their responses to different environments relative to the choice of testers.

Matzinger (1953) showed that as the heterogeneity of the testers increased, the component of variance of the line by tester interaction decreased. Comparisons of estimated components of variance for line by tester interactions support these conclusions for related line testers (Table 8.4).

Estimates of the line by tester interaction for the two related inbred lines were about twice those for the two broad genetic base testers. Also, line by tester interactions of S8 lines were about three times greater than for the S1 lines for both groups of testers. However, the line by tester interaction for the two elite line testers (B73 and Mo17) was similar for the S1 and S8 lines. For the 50 unselected lines included in testcrosses, line by tester interactions were greater for narrow genetic base testers, but interaction components of variance were smaller than testcross components of
Table 8.4  Line by tester interaction component of variance for three groups of testers

<table>
<thead>
<tr>
<th>Testers</th>
<th>Line × tester interaction</th>
<th>S₁</th>
<th>S₈</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS and BS13(S)C1</td>
<td></td>
<td>4.2 ± 3.0</td>
<td>12.3 ± 4.8</td>
</tr>
<tr>
<td>BSSS-222 and B73</td>
<td></td>
<td>7.3 ± 3.6</td>
<td>20.8 ± 6.6</td>
</tr>
<tr>
<td>B73 and Mo17</td>
<td></td>
<td>8.7 ± 4.4</td>
<td>6.8 ± 4.0</td>
</tr>
<tr>
<td>Inbred</td>
<td></td>
<td></td>
<td>17.2</td>
</tr>
<tr>
<td>Single cross</td>
<td></td>
<td></td>
<td>11.9</td>
</tr>
<tr>
<td>Double cross</td>
<td></td>
<td></td>
<td>6.5</td>
</tr>
</tbody>
</table>

Therefore, it seems that narrow genetic base testers can be effectively used to identify lines having good GCA. The data for an unselected group of lines support those reported by Sprague and Tatum (1942). Additional data of the study were reported by Lopez-Perez (1979) and Hallauer and Lopez-Perez (1979). Most of the evidence, though, was in agreement with Hull’s (1945) hypothesis that the most efficient tester is one having a lower frequency of favorable alleles.

8.1.3 Expected Gain from Selection

Allison and Curnow (1966) studied the effect of different testers on expected progress after selection among testcrosses. It was shown in Chapter 2 that expected change in the population mean is

\[ \Delta G = \left(\frac{k}{\hat{\sigma}}\right)p(1 - p)\left[a + (1 - 2p)d\right] \left[a + (1 - 2r)d\right] \]

This expression can be rewritten as

\[ \Delta G = \left(\frac{k}{2\hat{\sigma}}\right)\left[\hat{\sigma}_A^2 - 4p(1 - p)\alpha (r - p) d\right] \]

where \( \alpha = [a + (1 - 2p)d] \) is the average effect of substitution of the favorable allele and \( \hat{\sigma}_A^2 = 2p(1 - p)\alpha^2 \) is the additive genetic variance in the non-inbred parental population.

Assuming \( \hat{\sigma}_A^2 > 0 \), changes in population means using various testers are as shown in Table 8.5.

It is clear that under no dominance there is no difference among testers. With partial to complete positive dominance \((d > 0)\), the best tester is homozygous recessive. Under-dominance is not expected to be common, but if it occurs the best tester would be the homozygous dominant. The change in mean \( \Delta G = 0 \) for either \( p = 0.0 \) or \( p = 1.0 \) for any degree of dominance. However, with overdominance there is an optimum gene frequency \( p_0 \) \((0 < p_0 < 1)\), for which \( \Delta p = 0 \). This point of equilibrium occurs when \( p_0 = 0.5(1 + a/d) \), so an increase in the population mean occurs when \( p < p_0 \) and selection is for high-yielding crosses or when \( p > p_0 \) and
selection is for low-yielding crosses. In all instances the best tester seems to be the recessive homozygote. The probability of an inbred line being homozygous recessive for all loci is infinitesimally small. Therefore, the problem of tester choice is to find an inbred line homozygous recessive at the most important loci or a variety with low gene frequency at the most important loci. Despite the gene frequency in the tester, the importance of a locus is related to its contribution to the phenotype because a small increase in gene frequency would result in a relatively large contribution toward changing the phenotypic mean. However, important loci in the tester are those whose contributions to increase the mean are of such magnitude to be significantly influenced by the gene frequency in the tester.

The challenge of tester choice cannot be studied in individual genes because they act together in the control of a quantitative trait. From the theoretical point of view some comments can enhance understanding of the problem. The formula for the expected change in phenotypic mean can be rewritten as

\[ \Delta G = \left( \frac{k}{\hat{\sigma}} \right) p (1-p) \left[ a^2 - 2p(1-2p)d^2 \right] + \left( 2 \frac{k}{\hat{\sigma}} \right) p (1-p) \left[ ad + (1-2p)d^2 \right] s \]

Thus, the component of \( \Delta G \) that depends on gene frequency in the tester is

\[ (2k/\hat{\sigma})p(1-p) \left[ ad + (1-2p)d^2 \right] s, \]

and the most important locus is that which exhibits a maximum value for this component. Assuming all genes have the same effect, the maximum coefficient for \( s = 1 - r \) will depend on gene frequency and degree of dominance (Table 8.6).

For example, at a degree of dominance of 0.5, for a given gene frequency in the tester and a given level of additive effect, maximum change in the mean as affected by the tester would be due to an initial gene frequency \( p = 0.392 \). Under the same conditions, a gene at frequency, say, 0.8 or 0.2 would be less important in the tester. In this case the change in mean could be greater but the portion of these changes that depends on tester gene frequency would be smaller. For no dominance \( (d = 0) \) all values equal zero in Table 8.6. For positive dominance \( (d > 0) \) the relative importance of a gene in the tester increases with the degree of dominance for gene frequencies \( p \leq 0.5 \), but there is a maximum for gene frequencies \( p > 0.5 \) in the population.

---

### Table 8.5 Expected genetic gain at one locus level in a population for three types of testers

<table>
<thead>
<tr>
<th>Tester</th>
<th>Gene frequency</th>
<th>Expected change in population meana</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homozygous dominant</td>
<td>( r = 1 )</td>
<td>( \frac{k}{2\hat{\sigma}} \left[ \hat{\sigma}^2_A - 4pq^2 , ad \right] )</td>
</tr>
<tr>
<td>Homozygous recessive</td>
<td>( r = 0 )</td>
<td>( \frac{k}{2\hat{\sigma}} \left[ \hat{\sigma}^2_A + 4p^2q , ad \right] )</td>
</tr>
<tr>
<td>Parental population</td>
<td>( r = p )</td>
<td>( \frac{k}{2\hat{\sigma}} \left[ \hat{\sigma}^2_A \right] )</td>
</tr>
</tbody>
</table>

aAdapted from Allison and Curnow (1966), where \( \hat{\sigma}^2_A \) is the additive genetic variance at the locus and \( \alpha = a + (q - p)d \) is the average effect of substitution of the favorable allele.
Table 8.6 Coefficients for $s$ (frequency of recessive allele in the tester) in $\Delta G$ as a function of gene frequency in the population ($p$) and degree of dominance ($d/a$) for a fixed value of $(2ka^2/\sigma^2)$

<table>
<thead>
<tr>
<th>Allele frequency ($p$)</th>
<th>0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0</td>
<td>0.021</td>
<td>0.048</td>
<td>0.063</td>
<td>0.080</td>
<td>0.118</td>
<td>0.162</td>
<td>0.212</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0</td>
<td>0.036</td>
<td>0.079</td>
<td>0.104</td>
<td>0.131</td>
<td>0.189</td>
<td>0.256</td>
<td>0.330</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0</td>
<td>0.045</td>
<td>0.097</td>
<td>0.126</td>
<td>0.156</td>
<td>0.222</td>
<td>0.294</td>
<td>0.373</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0</td>
<td>0.050</td>
<td>0.104</td>
<td>0.132</td>
<td>0.161</td>
<td>0.223</td>
<td>0.288</td>
<td>0.357</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0</td>
<td>0.050</td>
<td>0.100</td>
<td>0.125</td>
<td>0.150</td>
<td>0.200</td>
<td>0.250</td>
<td>0.300</td>
</tr>
<tr>
<td>0.6</td>
<td>0.0</td>
<td>0.046</td>
<td>0.088</td>
<td>0.108</td>
<td>0.127</td>
<td>0.161</td>
<td>0.192</td>
<td>0.219</td>
</tr>
<tr>
<td>0.7</td>
<td>0.0</td>
<td>0.039</td>
<td>0.071</td>
<td>0.084</td>
<td>0.096</td>
<td>0.114</td>
<td>0.126</td>
<td>0.131</td>
</tr>
<tr>
<td>0.8</td>
<td>0.0</td>
<td>0.028</td>
<td>0.049</td>
<td>0.056</td>
<td>0.061</td>
<td>0.067</td>
<td>0.064</td>
<td>0.054</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0</td>
<td>0.015</td>
<td>0.024</td>
<td>0.027</td>
<td>0.028</td>
<td>0.026</td>
<td>0.018</td>
<td>0.004</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$p$ (at maximum)</td>
<td>—</td>
<td>0.451</td>
<td>0.410</td>
<td>0.392</td>
<td>0.377</td>
<td>0.352</td>
<td>0.333</td>
<td>0.319</td>
</tr>
</tbody>
</table>

Allison and Curnow’s (1966) general results led to the conclusion that if there is no overdominance any tester leads to an improvement in the population mean except for the case of complete dominance and genes fixed at frequency $p = 1$ at all loci in the tester. However, with overdominance the choice of a more suitable tester may lead to a decrease in mean yield unless the direction of selection is changed. Low gene frequency for the most important loci was shown to be desirable in the tester. The choice of the parental variety as tester was thought to assure effective selection since only in this case would a high frequency of the recessive allele in the tester always be associated with a high frequency of the recessive allele in the material being tested. An unrelated, low-performance variety would be a good tester only if the low performance is due to a low frequency of genes at important loci.

Cress (1966) emphasized that the choice of a tester to maximize gain from selection for a heterogeneous population depends on the average performance of the testcrosses. This meant the tester with the highest average cross performance is chosen. Nevertheless, ‘unless the selected individuals are to be used immediately in hybrid combination with the tester, this emphasis on heterotic response is misplaced. The heterotic response reveals little concerning the genetic potential and nothing concerning the expected rate of progress from selection.’ Therefore, it is clear from his statement that choice of tester could be dependent on the stage of breeding, e.g., pre-breeding for maximizing genetic improvement of heterogeneous populations vs. early- and late-generation hybrid testing for inbred line development.

Evidence presented by Hull (1945), Rawlings and Thompson (1962), Comstock (1964), and Allison and Curnow (1966) led to similar conclusions with respect to the choice of tester:
8.1 Theory

1) An inbred line homozygous recessive at important loci for discriminating among inbred lines to identify hybrids.
2) A population with low gene frequency at important loci for population improvement in a recurrent selection scheme.

8.1.4 Reciprocal Recurrent Selection

In a reciprocal recurrent selection program there is no opportunity for the choice of testers because each population is the natural tester of its counterpart. However, results reported by Darrah et al. (1972) and Horner et al. (1973) showed that the genetic variance among testcrosses was about twice as large when inbred testers were used as when a non-inbred population tester was used in intra-population recurrent selection programs. These findings led Russell and Eberhart (1975) to propose that progress in RRS would be greater if inbred lines extracted from the parental populations were used as testers instead of the populations themselves, as in the original procedure proposed by Comstock et al. (1949). Comstock (1977) further analyzed, theoretically, the possible implications of such a modification. Theory based on a model that assumes no multiple alleles, no epistasis, and linkage equilibrium provides the expressions shown in Table 8.7 for variance among testcrosses and expected change in gene frequency at the one-locus level.

Table 8.7 Genetic variance and expected change in gene frequency for three different testers

<table>
<thead>
<tr>
<th>Tester</th>
<th>Genetic variance</th>
<th>Expected change in gene frequency ([E(\Delta p)])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opposite population</td>
<td>(V_p = \left(\frac{1}{2}\right)p(1-p)[a+(1-2r)d]^2)</td>
<td>(\frac{k}{2\sigma p} p (1-p) [a + (1 - 2r) d ])</td>
</tr>
<tr>
<td>Inbred homozygous for better allele</td>
<td>(V_1 = \left(\frac{1}{2}\right)p(1-p)[a-d]^2)</td>
<td>(\frac{k}{2\sigma p} p (1-p) [a - d ] = \Delta_1)</td>
</tr>
<tr>
<td>Inbred homozygous for poorer allele</td>
<td>(V_0 = \left(\frac{1}{2}\right)p(1-p)[a+d]^2)</td>
<td>(\frac{k}{2\sigma p} p (1-p) [a + d ] = \Delta_2)</td>
</tr>
</tbody>
</table>

Adapted from Comstock (1977)

If inbred lines are extracted from the tester population, there is a probability \(r\) that they are homozygous for the better allele AA and a probability \(1 - r\) that they are homozygous for the poorer allele aa. Thus expected genetic variance among testcrosses using an inbred line as tester is

\[
V_I = rV_1 + (1-r) V_0 = \left(\frac{1}{2}\right)p (1-p) [a^2 + 2(1-2r)ad + d^2]
\]

Therefore,

\[
V_I = 2Vp \text{ when } a^2 + 2(1-2r)ad + d^2 = 2[a + (1 - 2r)d]^2
\]

when, for instance, \(r = (1 + c) \pm \sqrt{2c^2 - (\frac{1}{4})c}\), where \(c = d/a\) is a measure of the degree of dominance (equivalent to \(a\) in Comstock’s notation).
A solution for $r$ is possible only for $c \geq 0.5$, as shown by Comstock (1977). On the other hand, the expected change in gene frequency using inbreds from the tester population is

$$E(\Delta p) = r \Delta_1 + (1 - r) \Delta_2 = \left[ \frac{k}{2\hat{\sigma}} \right] p (1 - p) \left[ a + (1 - 2r) d \right]$$

which equals expected change in gene frequency when using the opposite non-inbred population as tester. Such a comparison does not take into account differences in $k/(2\hat{\sigma})$ values; $k$ is a function of selection intensity and may be taken as a constant and $\hat{\sigma}$ is the phenotypic standard deviation of the testcross means, which is not expected to vary greatly mainly if the environment is the most important source of variation. Betran and Hallauer (1996) compared hybrids produced from unselected S7 inbred lines before and after nine cycles of reciprocal recurrent selection in BSSS and BSCB1. They found that hybrid performance significantly increased after nine cycles of selection, and the rate of improvement of the hybrids was similar to the rate of improvement of the population crosses.

Because inbred lines are extracted at random, it may occur that a particular sample gives greater changes in gene frequency because of an excess of recessive alleles over the expected number. In terms of expectation there is no theoretical reason to expect better results by the use of inbred lines rather than the population as tester. Comstock (1977) presented some evidence favoring use of a population as tester in an RRS program. First, phenotypic variance among testcrosses is smaller when the population is used as tester and genotype by environment interaction is also likely to be smaller. Second, variance of expected change in gene frequency $V(\Delta p)$ is smaller for the population as tester. As a consequence, the probability of poor allele fixation is also smaller. Comstock (1979) further extended the comparisons for multiple alleles to support his conclusions that, on average, rates of change in allele frequencies in RRS programs are not improved when inbred lines extracted from the RRS populations are used as testers rather than the populations themselves. Comstock (1979) also emphasized that when comparing testers to be employed in RRS, the critical parameter is expectation of allele frequency change per unit of time. The suggestion by Russell and Eberhart (1975) would be a consequence of the erroneous assumption that ‘because of the greater variance among the testcrosses with the inbred tester, gain from selection would be greater.’ This again agrees that choice of tester may vary depending on the breeding stage of a program (e.g., germplasm improvement or inbred line development).

### 8.1.5 Combining Ability in Testcrosses

The main difference between general and specific combining ability has been attributed to the tester genetic basis (broad or narrow genetic base). Such differences are essentially a matter of differences in gene frequency. In the broad base tester gene frequencies for different loci may vary from 0 to 1, whereas in the narrow base tester gene frequencies are limited to a few values. For an inbred line,
allele frequency at a locus may be 0 or 1 and in the single cross of homozygous lines allele frequencies may be 0, 0.5, or 1. In either case (broad or narrow base tester) selection can lead to change in the population mean as a result of selection for additive effects of genes. Therefore, when considering genotypes in a population or in other genetic materials that are to be evaluated in testcrosses, it seems difficult to distinguish between general and specific combining ability. Therefore, the expression combining ability should be used with a broader meaning.

The combining ability of genotypes in a population or in any other genetic materials can be estimated as shown in Table 8.8.

Genetically, combining ability is measured as

\[ c_i = \bar{T}_i - \bar{T} = (p_i - \bar{p}.) \left[ a + (1 - 2r) d \right] \]

where \( r \) is the frequency of the favorable allele in the tester. Combining ability \( c_i \) is independent from dominance effects only for \( r = 0.5 \) or for \( d = 0 \) (no dominance). When gene frequency is fixed in the tester (\( r = 1.0 \)), \( c_i = (p_i - \bar{p}.) (a - d) \); and for complete dominance (\( d = a \)), \( c_i = 0 \) for every \( i \). In this case the tester is not able to discriminate among lines or populations being tested.

Variance among testcrosses \( \hat{\sigma}_t^2 \) (same as in Chapter 6) is an estimate of variance for combining ability \( \hat{\sigma}_c^2 \). Genetically, we have

\[ \hat{\sigma}_c^2 = \hat{\sigma}_p^2 \left[ a + (1 - 2r) d \right]^2 \]

It can be seen that \( \hat{\sigma}_c^2 \) depends on the square of the average effect of gene substitution in the tester and on the variance of average gene frequencies of the genetic materials under test. For the particular case of genotypes within a population, gene frequencies (for a favorable allele B) within genotypes are 1, 0.5, and 0 in BB, Bb, and bb, respectively. Because gene frequency has a binomial distribution, then

| Table 8.8 Combining ability of genotypes in a population or a set of populations (Y) in testcross with a population as tester |
|---|---|---|---|---|---|---|
| TC | G | F | \( f_g \) | P | F | \( f_y \) | Mean of TC | Combining ability (\( c_i \)) |
| 1 | BB | \( p^2 \) | 1 | \( P_1 \) | \( 1/n \) | \( p_1 \) | \( T_1 \) | \( T_1 - \bar{T} \) |
| 2 | Bb | \( 2pq \) | 0.5 | \( P_2 \) | \( 1/n \) | \( p_2 \) | \( T_2 \) | \( T_2 - \bar{T} \) |
| 3 | bb | \( q^2 \) | 0 | \( P_3 \) | \( 1/n \) | \( p_3 \) | \( T_3 \) | \( T_3 - \bar{T} \) |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| \( N \) | — | — | — | \( P_n \) | \( 1/n \) | \( p_n \) | \( T_n \) | \( T_n - \bar{T} \) |
| Avg. | — | — | \( P \) | — | \( p \) | \( \bar{p} \) | \( \bar{T} \) | \( \bar{T} \) |

\[^a\text{Populations with genotypes in Hardy–Weinberg proportion; } f_g \text{ and } f_y \text{ are the average gene frequency within genotypes and within populations, respectively.} \]

TC testcross, G genotype, F frequency, Avg. average
\[
\hat{\sigma}_p^2 = p(1-p)/2
\]

and

\[
\hat{\sigma}_c^2 = [p(1-p)/2][a + (1-2r)d]^2
\]

which is the variance among testcrosses (see Chapter 2). If the genotypes represent a random sample of inbred lines \((F = 1)\), the genotypic array is \(p(BB):(1-p)(bb)\) with gene frequencies 1 and 0 within genotypes, respectively. Thus,

\[
\hat{\sigma}_c^2 = p(1-p)[a + (1-2r)d]^2
\]

for a random sample of inbred lines

In general, for any level of inbreeding \(F\) in the base population the variance for combining ability is

\[
\hat{\sigma}_c^2 = [p(1-p)/2](1+F)[a + (1-2r)d]^2
\]

The variance \(\hat{\sigma}_c^2\) must be extended to summation over all loci, and when inbred lines are used as testers the variance will depend on the balance of the number of loci with frequency 1 or 0. An excess of important loci with frequency 0 will always be advantageous relative to increasing the variance and changing the population mean through selection.

All the theoretical aspects so far presented deal with simple models at one locus level, so it is not possible to predict the relative magnitude of variances and expected changes in gene frequencies when results have to be extended to all loci controlling the character. Although genes controlling a quantitative trait cannot be handled individually (e.g., as QTL), it is reasonable to assume that they differ from one another with respect to magnitude and type of gene action and that complex interactions exist among them. It is also reasonable to suppose that alleles occur with different frequencies within populations. Most genes at heterogeneous populations are at intermediate frequencies and few genes have frequencies near the fixation (0 or 1). Therefore, additional theory for the study of problems (such as that presented in this section) considering the whole population of genes and their properties would be straightforward and should provide additional important tools for the use of quantitative genetic theory in maize breeding.

Additional studies related to the choice of testers have been conducted to determine

(1) relation of line performance to hybrid performance,
(2) effectiveness of visual selection for developing lines that have good combining ability,
(3) genetic diversity of the lines being tested,
(4) appropriate stage of inbred line development for testing for combining ability, and
(5) relative importance of general and specific combining abilities.
Any one study, however, does not provide information on all the above five items. Data are often conflicting and interpretations differ. In some instances relative merits of the art and science of maize breeding are in conflict.

The development of cultivars in self-pollinating crops (e.g., pure lines) differs from the development of cultivars in cross-pollinating crops (e.g., hybrids). The development of new maize lines is clearly easier than determination of their worth in hybrids. Maize lines can be excellent for several traits but unless they are also excellent in hybrid combinations they may not make it commercially. All lines are not equally good in hybrids and methods are needed to screen for elite lines for potential use in hybrids. During inbred line development it is a common practice to discard lines that have obvious morphological deformities, are difficult to maintain because of either poor pollen production or ear shoot development, are obviously susceptible to diseases and insects, and have poor plant vigor. Phenotypic elimination of lines is effective for many traits for use in hybrids, but effective phenotypic elimination of lines that have poor combining ability has not been resolved. Some traits, such as disease and insect resistance, are used in discarding lines because they are not acceptable for commercial use although they may have good yield potential in hybrids in a pest-free environment.

There is a wide variety of methods to develop maize inbred lines. The term pedigree selection in maize is thought to mean using only F2 and BC populations involving elite lines and their hybrid testing. This usage is not valid for maize since pedigree selection methods were used to develop B73 (following half-sib recurrent selection in BSSS) and many other successful lines to produce commercial hybrids. Recurrent selection procedures are applied to genetically broad-based populations not only to increase the frequencies of favorable alleles but also to increase the efficiency of pedigree selection. Pedigree breeding methods are often used in the public sector to derive lines from germplasm improved by recurrent selection (see Fig. 6.1).

### 8.2 Correlations Between Lines and Hybrids

Relations of inbred line traits to the same traits in hybrids have been one of the major limitations of inbred and hybrid development programs. Because yield trials are expensive to conduct, any information on inbred lines that is indicative of their performance in hybrids (e.g., correlations, heterosis prediction) is desirable in order to eliminate the need for making crosses and conducting an extensive amount of yield trials. Hence, it is desirable to investigate possible methods to reduce testing of inbred lines in hybrids and to determine if expression of traits in inbred lines is transmissible to their hybrids.

Correlation studies between inbred line traits and either the same or different traits in their hybrids have been used to determine the effectiveness of selection on hybrid performance (Table 8.9). These correlations are independent of those in Chapter 5, which were obtained from progenies developed by the use of mating
designs. The correlations in Table 8.9 were, with one exception, obtained from what was considered the best and most vigorous inbred lines that had survived selection and their crosses, mostly single crosses. Consequently, sample size usually was small and the correlations cannot be interpreted relative to some population. Simple phenotypic correlations were calculated in most instances.

Kiesselbach (1922), Richey (1924), and Richey and Mayer (1925) reported evidence that certain inbred lines were better than others in transmitting high-yielding ability to their crosses. Although there was a general tendency for some inbred lines to be better than others in crosses, Richey and Mayer (1925) emphasized that the lack of any definitive correlation between yields of parent inbred lines and their crosses indicates that selection for combining ability in the final analysis must be based on the performance of the lines in crosses rather than on the inbred lines per se. Richey (1924) also emphasized that the best chances of success seem to be in the use of large numbers and in avoiding pre-conceived notions as to what are the best lines.

Jorgenson and Brewbaker (1927), Nilsson-Leissner (1927), Jenkins (1929), Johnson and Hayes (1936), and Hayes and Johnson (1939) studied the relation of traits of inbred lines to the same traits in their crosses and to yield of their crosses. In most instances simple correlations were obtained that were considered of sufficient magnitude to be useful in selection, but Johnson and Hayes (1936) concluded that yield of inbred lines was not significantly correlated with combining ability of testcrosses. Relatively large multiple correlations between yield of crosses with yield and other traits of inbred lines were obtained by Jorgenson and Brewbaker (1927), Nilsson-Leissner (1927), and Hayes and Johnson (1939). In all instances $R$ values, ranging from 0.61 to 0.82, indicated that more than 50% of the total variation in yield of the crosses was accounted for by inbred line traits.

Jenkins (1929), in a comprehensive study for the correlation of inbred traits with the same traits in their hybrids, determined the correlations between

1. various traits of parental inbred lines and the same traits of the crosses and
2. traits of parental inbred lines and means of the same traits of all crosses that included the inbred line common in all crosses.

Positive correlations were found for all 19 traits studied but they were small in most instances. In the first case, none of the characters of the parental inbreds was closely related to its $F_1$ cross; correlations ranged from $-0.10$ to 0.24. Correlations between yield of parental inbred lines and yield of their $F_1$ crosses were 0.14 for case 1 and 0.20 for case 2 (Table 8.9). Multiple correlations for different combinations of inbred line traits and yield of their $F_1$ crosses ranged from 0.20 to 0.42. In the second case, correlations were generally larger, ranging from 0.25 to 0.67. Because of the positive correlations between traits of inbred lines and their crosses, Jenkins (1929) concluded that the correlations were, in some instances, of predictive value. Gama and Hallauer (1977) reported a similar study for unselected inbred lines and their single-cross hybrids developed from Iowa Stiff Stalk Synthetic. Genetic correlations were determined for eight traits and they were too small in all instances to have
Table 8.9  Summary of six studies reporting the relation between traits of inbred lines and the same traits of their crosses

<table>
<thead>
<tr>
<th>Trait</th>
<th>Jenkins (1929)</th>
<th>Gama and Hallauer (1977)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Yield</td>
<td>0.46 ± 0.05</td>
<td>0.50 ± 0.08</td>
</tr>
<tr>
<td>Ear length</td>
<td>0.73 ± 0.02</td>
<td>0.58 ± 0.08</td>
</tr>
<tr>
<td>Ear diameter</td>
<td>0.92 ± 0.01</td>
<td>0.63 ± 0.06</td>
</tr>
<tr>
<td>Kernel row no.</td>
<td>0.95 ± 0.01</td>
<td>0.79 ± 0.04</td>
</tr>
<tr>
<td>Kernel depth</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Plant height</td>
<td>0.75 ± 0.02</td>
<td>0.48 ± 0.08</td>
</tr>
<tr>
<td>Days to flower</td>
<td>−</td>
<td>0.29 ± 0.09</td>
</tr>
<tr>
<td>Prolificacy</td>
<td>0.48 ± 0.08</td>
<td>−</td>
</tr>
</tbody>
</table>

*aThe left-hand correlations are between one parental inbred line and the mean of all its crosses; the right-hand correlations are between the mean of the two parental inbreds and their specific cross; bEars from tillers included
predictive power. Correlations for yield, for example, were only 0.09 and 0.11 for cases 1 and 2, respectively. Multiple correlations of plant traits, ear traits, and plant and ear traits of inbred lines with yield of their crosses were also small.

Kyle and Stoneberg (1925), Hayes (1926), Jenkins (1929), Kovacs (1970), Obilana and Hallauer (1974), and Bartual and Hallauer (1976) determined correlations among traits of inbred lines; Kempton (1926), Jenkins (1929), El-Lakany and Russell (1971), and Silva (1974) determined correlations among traits of hybrids. Correlations by Hayes (1926) were for eight traits of inbred lines in different generations of inbreeding. Correlations of ear length and number of ears with yield were significantly positive. Jenkins (1929) and Kovacs (1970) determined correlations for several traits with yield of inbred lines in the same generation of inbreeding. Jenkins (1929) reported that yields of inbred lines were positively and significantly correlated with plant height, number of ears per plant, ear length, ear diameter, and shelling percentage, but they were negatively and significantly correlated with date of silking, ear shrinkage, chlorophyll grade, and ear shape index.

Kovacs (1970) reported data for five ear traits and yield for a selected group of inbred lines in a breeding study. Average correlations for ear traits of four groups of inbred lines with yield were 0.68 (ear length), 0.56 (number of kernel rows), 0.74 (number of kernels), 0.72 (kernel length), and 0.59 (thousand-seed weight). However, there were large differences in the correlations of lines that were considered good and poor for yielding ability. Obilana and Hallauer (1974) and Bartual and Hallauer (1976) evaluated two sets of unselected inbred lines derived from Iowa Stiff Stalk Synthetic and generated genetic correlations among 11 traits and the same traits with yield. Except for some of the correlations among ear components and ear components with yield, most correlations were small. Kernel depth (0.82 and 0.76) and kernels per row (0.86) had the greatest correlations with yield.

Additional correlation studies were given by Love (1912), Collins (1916), Love and Wentz (1917), Etheridge (1921), Kempton (1924, 1926), and Richey and Willier (1925). The studies calculated the correlations primarily to determine if some ear and plant traits could be used for yield improvement. Detailed data were taken on several different traits and positive and significant correlations usually were reported. Kempton (1926) reported significant correlations between length of tassel central spike and total ear length. The data were 0.27 (F2 plants) and 0.34 (F1 plants), both of which exceeded three times their standard errors. Similarly, significant correlations were obtained for number of tassel branches and number of ear kernel rows (0.19 and 0.17), but further analysis by partial correlations showed no genetic correlation between number of rows on the ear with either season or grain yield. All correlations were too small to adequately predict future selection progress and correlations would be restricted to the F2 population of a wide cross of the Mexican variety Jala and the popcorn variety Tom Thumb.

Love and Wentz (1917) summarized the literature and presented results from their studies for the traits listed on the early maize show cards (see Chapter 1) and yield. All correlations were small and not consistent among the different studies. Love and Wentz concluded the following:
The judge at a corn show or a farmer in selecting his seed corn cannot pick the high-yielding seed ears when judging from outward characters of the ears. It is evident that the points emphasized on a score card are of no value for seed ear purposes and are entirely for show purposes.

The only basis left for selecting high-yielding seed corn is the ear-to-row progeny test. These remarkable conclusions are equally valid today, although experimental techniques have been refined.

The relation between performance of inbred lines per se and their respective testcrosses is still challenging and difficult to predict. Jensen et al. (1983) conducted a study to simulate methods used in applied breeding programs. They developed lines from elite germplasm and evaluated the S₂ lines per se and their respective testcrosses at the S₂ and S₅ generations for grain yield. The correlation between S₂ lines per se and S₅ testcrosses was 0.14 compared with a correlation of 0.67 between S₂ and S₅ testcrosses. Jensen et al. (1983) concluded that S₂ testcrosses were a better predictor of S₅ testcrosses than S₂ lines per se for grain yield. Smith (1986) conducted a computer simulation study to compare performance of lines per se with their respective testcross performance. He used a genetic model that included a large number of loci with complete dominance effects and each of the lines was crossed with what was considered a good, average, and unrelated tester. Correlations between lines per se and their respective testcrosses were 0.22 (good tester), 0.28 (unrelated tester), and 0.34 (average tester). The effects of linkage and epistasis were not considered in determining the correlations but could increase predictive power (Dudley and Johnson, 2009). Smith’s (1986) conclusions were similar to those of Jensen et al. (1983) that line per se performance was not a good predictor of hybrid performance. Obaidi et al. (1998) also conducted a computer simulation study that emphasized line development but the selection criteria were based on testcross performance even as early as the S₀ generation.

Although several reported correlations were relatively large, nearly all the authors stated in their conclusions that comparative yield trials of the hybrids are needed. It is necessary, however, that selection be practiced during inbred line development. The ultimate use of inbred lines is in hybrids, but it is necessary that they possess certain standards of vigor and productiveness as a line before testing if the line is to have potential commercial value for hybrid seed production. In some cases it may be an advantage to the breeder if the inbred line traits are not correlated with yield of their crosses because selection against many traits can be emphasized without fear of discarding all the potentially high-yielding genotypes. If the correlation is small for traits that are not of economic importance (e.g., plant color), we would have a normal distribution on inbred lines with acceptable vigor and productiveness for testing in hybrids. For other traits (e.g., stalk lodging) we would benefit from a correlation of the trait between inbred lines and their hybrids because inbred lines with weak stalks could be discarded before testing. Jenkins (1929), for example, reported a correlation of 0.88 ± 0.03 for percentage of erect plants at harvest between inbred lines and their crosses, indicating that 77% of the variation in the crosses for erect plants could be associated with the inbred lines used in the
crosses. For commercial usage, this correlation is desirable in the inbred lines and their crosses.

Bauman (1981) conducted a survey among active maize breeders to determine traits they considered important in developing inbred lines, relative importance of traits, and how effective visual selection was for 17 different traits during inbreeding and selection within and among progenies. The effectiveness of visual selection during inbreeding was inversely related to the importance of the 17 traits \( r = -0.54^* \), as seen in Table 8.10. While some traits, based on breeder’s experience, have been relatively easier to improve genetically (e.g., flowering time and disease resistance), others have been challenging (e.g., grain yield, rate of dry down, drought tolerance, and root strength). Common sense of breeders seems to be related with the genetic complexity of traits. However, exceptions are possible as in the case of flowering time (Buckler et al., 2009) where it seems that the simplicity to achieve fast genetic gain in this trait by breeders might not be related to the complexity found in the number of minor QTL reported by the authors. As suggested by Obaidi et al.’s (1998) results, effective visual selection can be made for several traits, but the ultimate value of an inbred is determined by its hybrid performance. In summary it seems that effective selection can be made for certain traits, but the ultimate use of inbred lines in crosses must be determined from extensive yield evaluations of the crosses.

### Table 8.10 Effectiveness of visual selection among selected traits for selection among and within inbred progenies

<table>
<thead>
<tr>
<th>Trait</th>
<th>Importance</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain yield</td>
<td>1.1</td>
<td>3.2</td>
</tr>
<tr>
<td>Root strength</td>
<td>1.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Erect leaf habit</td>
<td>3.0</td>
<td>2.1</td>
</tr>
<tr>
<td>Flowering date</td>
<td>1.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Fast dry down</td>
<td>1.4</td>
<td>3.0</td>
</tr>
<tr>
<td>Drought tolerance</td>
<td>1.3</td>
<td>3.1</td>
</tr>
<tr>
<td>Grain quality</td>
<td>2.0</td>
<td>2.7</td>
</tr>
<tr>
<td>Cold tolerance</td>
<td>1.4</td>
<td>2.9</td>
</tr>
<tr>
<td>Disease resistance</td>
<td>2.4</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Relative importance: 1 = more important 4 = less important
Relative efficient: 1 = more effective 4 = less effective

### 8.3 Visual Selection

Studies determining the effectiveness of visual selection for inbred line development were reported by Jenkins (1935), Sprague and Miller (1952), Wellhausen and Wortman (1954), Osler et al. (1958), and Russell and Teich (1967). These studies included inbred lines that had undergone visual selection but comparisons were not possible between visual selection and no selection. Because some of the correlation studies showed a good relation between vigor and productiveness of the inbred
lines and their crosses, these investigators tested the effectiveness of visual selection compared to testcross evaluation.

Jenkins (1935) compared seven pairs of lines from the Iodent heterotic group and five pairs of lines derived from the Lancaster heterotic group in testcross evaluation trials for yield in generations S2 – S8, omitting S7. Each line in each generation was represented by a selected and a discarded sib strain. Yield tests of the Iodent lines showed that for the first two generations of inbreeding average yield of selected strains was significantly better than for discarded strains. For all generations of Lancaster lines and after the second generation of Iodent lines, testcross yields of the selected and discarded sib strains were not significantly different but the small differences consistently favored selected strains. Sprague and Miller (1952) also found no effect of visual selection on the combining ability of the selections during four generations of inbreeding for two sets of progenies.

Brown (1967) evaluated testcrosses of 20 lines visually selected as the most desirable phenotypes and 20 lines visually selected as the poorest phenotypes of 1,160 unselected lines from four open-pollinated varieties: Krug, Reid, Lancaster, and Midland. Yields were identical for the two groups of testcrosses (54.2 q/ha for selected vs. 54.2 q/ha for unselected). Brown (1967) concluded that ‘while it cannot be said that the type of selection practiced had any adverse effect on yield, neither did it contribute to any improvement in yield.’ However, testcrosses of visually selected lines, however, were superior to those of unselected lines for root and stalk lodging. Wellhausen and Wortman (1954) and Osler et al. (1958) found, however, that visual selection resulted in small positive gains in yield combining ability of the selections.

Russell and Teich (1967) compared inbred line performance and combining ability for yield of inbred lines developed by

1) visual selection within and among ear-to-row progenies during successive inbreeding generations and
2) selection within and among ear-to-row progenies based on testcross performance during successive inbreeding generations.

The inbred lines were developed from an F2 population of M14 × C103 under two plant densities:

a) High (59,304 plants/ha) and low (29,652 plants/ha) plant densities for visual selection.

b) High (59,304 plants/ha) and low (39,536 plants/ha) plant densities for testcross selection.

Testcross evaluations were made in the F2, F3, and F4 generations, whereas visual selection was practiced in the F2 through the F6 generations. Inbred selections for both selection schemes and for both plant densities, M14 and C103 inbred lines, and M14 × C103, were crossed with WF9 × I205 (tester for the testcross selection) and IA4810 (double-cross hybrid chosen to give a measure of GCA). The inbred lines
themselves were also evaluated. For simplification, the groups of materials were designated as follows:

- **Group 0**: Inbred lines selected by testcross performance at both rates
- **Groups 1, 2**: Inbred lines selected by testcross performance in low (1) and high (2) rates
- **Groups 3, 4**: Inbred lines visually selected in low (3) and high (4) rates
- **Groups 5, 6, 7**: Testcrosses of M14 (5), C103 (6), and M14 × C103 (7)

A comparison of group means showed that selection as measured by combining ability was effective for groups 0, 1, 2, and 4 but not for group 3 (visual selection at low density) when compared with (M14 × C103) × tester. Group 0 lines were superior to groups 1, 2, 3, and 4 lines, and the lines in group 0 showed less interaction with environments. Although Russell and Teich (1967) concluded that visual selection of inbred line performance in high stands was as effective as selection by extensive testcrossing, it seems that more extensive evaluation of the lines in group 0 resulted in higher yielding genotypes with greater stability over environments. Because testcross evaluations were at two plant densities, additional information on environmental response was available for the lines in group 0. In addition, evaluating testcrosses at low densities could be an option to increase the number of plots. Inbred line yields of groups 3 and 4 were superior to yields of groups 1 and 2, and yields of inbred lines selected at high density were superior to those selected at low density. However, the lack of genotype × density (and other type of interactions) does not justify duplicating selections across environments (Carena and Cross, 2003; Hyrkas and Carena, 2005; Carena et al., 2009).

Visual selection is always practiced by the maize breeder during each generation of inbreeding for inbred line development. Bauman’s (1981) survey showed that visual selection was considered effective for several traits (e.g., date of flowering, plant color, and plant stature) during line development, but for other important traits (e.g., grain yield and root strength) visual selection had minimal value for how the lines may perform in crosses. Effective visual selection among and within progenies for the expression of yield in crosses remains questionable. The evidence shows, however, that visual selection for several traits during line development is not detrimental in crosses for yield.

Attempting to measure effectiveness of visual selection in inbred line development is difficult because of the phenomenon of sampling. In some sets of selected materials, a relation may be obtained between some traits of the inbred lines and expression for yield in crosses, whereas in others there may be no relation. The differences may be due to sampling and the original assemblage of genes in the population. For an unselected set of inbred lines, Gama and Hallauer (1977) failed to detect any correlations of predictive value.

It seems that the art of plant breeding, or visual selection, has a place in maize breeding. Because vigor and productiveness of inbred lines are important for the cost-effective production of commercial single crosses, it is imperative that inbred lines whose traits do not impair the final products (high-performance hybrids) be easily reproducible to plant on sizable areas. Visual selection in modifying traits of
the lines themselves is generally accepted, but how this influences combining ability remains to be investigated.

### 8.4 Genetic Diversity

The need of genetic diversity between inbred lines used in crosses is generally accepted. In the past, general experience usually but not always showed that crosses of unrelated genotypes contributed to greater yields. It usually required extensive evaluation trials to determine the unique combination of two parental inbred lines. Although the genetic basis of the importance of genetic diversity was not clear, general experience showed that the better hybrids involved inbred lines derived from two or more genetic backgrounds (Hayes, 1963).

Empirical data on the effects of genetic diversity of parental inbred lines on hybrid performance were given by Hayes and Johnson (1939), Wu (1939), Eckhardt and Bryan (1940a,b), Johnson and Hayes (1940), Cowan (1943), and Griffing (1953). In all instances parental inbred lines of related origin consistently produced lower yielding crosses than those that had one or no parent lines in common. There were some exceptions, but the greater frequency of high-yielding crosses involved lines from different varieties. Most of the varieties were of US Corn Belt origin, but Griffing (1953) showed that parental lines including germplasm from outside the US Corn Belt also contributed to the crosses of diverse origin.

The importance of maintaining genetic diversity of the parent inbred lines was emphasized by the North Central Corn Breeding Research Committee (NCR-2, later renamed as NCR-167 and currently as NCCC167) in 1939 when it developed the A and B grouping of inbred lines. The designation of inbred lines in the A and B grouping was somewhat arbitrary, but it was an attempt to minimize gene exchange between the two gene pools. Strict enforcement of this arbitrary grouping was impossible, but it was recognized that opportunities for the production of high-yielding hybrids would be enhanced by crossing and testing inbred lines from the two groups rather than within the two groups. At that time, genetic erosion of the two groups was a concern because of pedigree selection in F2 populations of the best single-cross hybrids. As a consequence, this was the cause of the increasing interest for introducing exotic germplasm in breeding programs. In the US Corn Belt inbred lines of Reid origin crossed with those of Lancaster origin produced higher yielding crosses on the average. Similarly, crosses of lines originating from flint and dent varieties were commonly used in Europe and other areas of the world. Unlike correlation and visual selection, however, genetic diversity of the lines used in crosses was generally recognized to be important.

Crosses of more distantly related parents have shown greater heterosis than crosses of more closely related parents (Brown, 1950; Anderson and Brown, 1952; Moll et al., 1965; Melchinger, 1999). The establishment of heterotic patterns among varieties and lines had important diversity implications for developing new inbred lines as potential seed stocks in hybrids. Heterotic patterns are crosses between
known genotypes expressing a high level of heterosis (Carena and Hallauer, 2001). They played an important role to separate genotypes based on pedigree and heterotic relationships. Since the development of maize inbred lines is limited to their inclusion into specific diversity groups of lines with common pedigree or relationship (e.g., heterotic groups) it was thought that the use of molecular markers would increase the efficiency of heterosis prediction. The use of DNA marker data has been useful to complement pedigree information and assign diverse lines into heterotic groups (Melchinger et al., 1992; Mumm and Dudley, 1994; Smith et al., 1997; Senior et al., 1998; Barata and Carena, 2006; Sonnino et al., 2007) but of limiting usefulness for predicting good heterotic combinations. Therefore, evaluating the performance of crosses among groups based upon genetically diverse parents has been considered essential to identify promising heterotic patterns (Melchinger, 1999). Utilizing data from not only molecular information (e.g., marker data) but also from yield trials (e.g., testcrosses and mating designs) is an alternative (Barata and Carena, 2006) as long as there is a good methodology to link both types of data.

8.5 Testing Stage

The stage of testing inbred lines for combining ability has received special attention. Inbreeding five to seven generations accompanied by visual selection usually was completed before attention was given to the performance of the inbred lines in crosses (Bauman, 1981). Not until the publication of Jenkins’ study (1935) was empirical evidence given comparing the testcrosses of lines in different generations of inbreeding. Richey and Mayer (1925) concluded from their study that comparison of crosses indicates no general advantage for crosses made after five generations of inbreeding over analogous crosses made after three generations of inbreeding. This suggested to them that there is little inherent relation between the yield of a cross and the number of generations of inbreeding of its parent lines before crossing.

Because correlation studies of inbred line traits with the same traits of their F1 crosses (and hence visual selection during the inbreeding generations) did not seem to provide a satisfactory index of the yield potential of inbred lines in crosses, the next logical step was to determine if cross performance in the early generations of inbreeding was predictive of their performance at increased levels of inbreeding. It is generally acknowledged that isolation of genotypes from a breeding population depends on sufficiency of sampling. Because the potential ceiling of any derived line is determined at the time of the first selfing generation of an S0 plant, it seemed logical that better sampling techniques could be used if material could be discarded in early generations of inbreeding. Thus greater attention and effort could be expended on selection within progenies of the genotypes saved on the basis of the preliminary early yield test information.

Jenkins (1935), in the study evaluating effectiveness of visual selection, reported data that supported early testing as an effective procedure for saving lines that have greater potential than others. Comparisons of selected and discarded sib strains for the seven lines of the Iodent series and the five lines of the Lancaster series showed that association with generations was not significant for the Iodent series and was
not significant after the second generation of inbreeding for the Lancaster series. On the basis of his results, Jenkins concluded that ‘the inbred lines acquired their individuality as parents of testcrosses very early in the inbreeding process and remained relatively stable thereafter.’ His conclusion stimulated thinking about the merits of early testing to determine which progenies should be maintained in the nursery for further selection and inbreeding. In some respects, Jenkins’ conclusion seemed reasonable because small populations of each progeny usually are included in the breeding nursery; low heritability of yield, linkage of genes, and limited effectiveness of visual selection would be all factors influencing to a point that the selection for more desirable gene combinations for yield would be small. The probabilities of success would decrease with each generation of inbreeding.

Sprague (1946) reported the results of a program designed to compare the relative merits of early testing, as suggested by Jenkins (1935) and Sprague (1939). Sprague (1946) selected 167 phenotypically desirable S₀ plants of Iowa Stiff Stalk Synthetic, selfed each S₀ plant, and testcrossed to the double-cross tester IA13. Yields of the 167 testcrosses were normally distributed and ranged from 38.6 to 63.0 q/ha. The wide range of the selected S₀ plants for yield also is evidence of the poor relation between visual selection and yield of crosses. The difference (6.0 q/ha) necessary for significance at the 5% level showed that four of the crosses were significantly poorer than Iowa Stiff Stalk Synthetic and two were significantly better than the double-cross IA13. Two samples based on the testcross information were subsequently studied:

1. S₁ lines representing the best 10% of the 167 testcrosses were self-pollinated and crossed with IA13, the tester.
2. A group of 12 lines was chosen that represented a seriated sampling of the 167 testcrosses in which 20 self-pollinations and testcrosses were attempted with each family.

Only six families ultimately were available for testing in the second group. Distribution of S₁ testcross yields clearly indicated that S₀ plants exhibiting high combining ability transmitted this trait to their S₁ progeny. Significant differences were noted among the 20 testcrosses within each of the six families, which suggested that heterozygosity was similar for S₀ plants having good and poor testcross yields. Distributions from the four best yielding families were not significantly different, but they were significantly different from the two poorest yielding families. Sprague (1946) compared three of the selected samples in the S₃ generations in testcrosses with five standard lines. All possible crosses were produced and the three selected lines were superior to the standard lines for yield and resistance to root and stalk lodging. Sprague (1946) concluded from these data that early testing might be a useful tool in a breeding program and the data seem to support Jenkins (1935), Johnson and Hayes (1940), Cowan (1943), and Green (1948a,b) that combining ability is a heritable trait.

Lonnquist (1950) conducted an early testing study for a series of selected plants from a strain of Krug Yellow Dent. After the S₀ plant testcross performance data
were available, Lonnquist practiced divergent selection for good and poor testcross yields for a group of lines that had the best and poorest S₀ testcross yields. Results from the first four generations of inbreeding showed that testcross combining ability could be modified by a combination of selection and testing. Divergent selection for good and poor combining ability for three generations within the original lines that had poor combining ability produced S₄ lines that were not significantly different from those in the good combining group selected for three generations for poor combining ability. Thus selection in the low combining group was not fruitful and additional selection and testing should be concentrated on the progenies that exhibit the best combining ability in the S₁ generation.

Richey (1945), Singleton and Nelson (1945), and Payne and Hayes (1949) expressed doubt about the value of early testing for combining ability. Richey’s conclusions were based on a re-analysis of Jenkins’ data indicating that his interpretations for the value of early testing were not warranted. Richey’s conclusions are somewhat surprising in view of the conclusions given by Richey and Mayer in 1925. Singleton and Nelson’s data were for a group of selected lines continued for three generations of inbreeding; they concluded that early testing was ineffective, but there were no significant differences among the 10 lines studied. Payne and Hayes presented data comparing the S₀ and S₁ testcrosses and concluded that testing of S₀ plants was of doubtful value. However, as Sprague (1955) has pointed out, Payne and Hayes’ data indicated that the S₀ plants with the best combining ability produced a greater percentage of good combining S₁ crosses than the S₀ plants with poor combining ability.

Lopez-Perez (1979) also was able to compare yields of the testcrosses of the same lines at the S₁ and S₈ levels of inbreeding (Hallauer and Lopez-Perez, 1979). Genetic correlations were determined between the S₁ and S₈ testcrosses, among testers at the S₁ and S₈ levels, and between S₈ testcrosses and yields of S₇ lines per se (see Table 8.11).

### Table 8.11 Genetic correlations between S₁ and S₈ testcrosses and between S₇ lines per se and S₈ testcrosses for five testers for yield

<table>
<thead>
<tr>
<th>Testers</th>
<th>BSSS</th>
<th>BS13(S)C1</th>
<th>BSSS-222</th>
<th>B73</th>
<th>Mo17</th>
<th>Avg.</th>
<th>S₇ linesᵃ</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSS</td>
<td>0.20ᵇ</td>
<td>0.74</td>
<td>0.25</td>
<td>0.22</td>
<td>0.65</td>
<td>0.46</td>
<td>0.17</td>
</tr>
<tr>
<td>BS13(S)C1</td>
<td>0.68</td>
<td>0.17</td>
<td>0.71</td>
<td>0.48</td>
<td>0.91</td>
<td>0.71</td>
<td>0.10</td>
</tr>
<tr>
<td>BSSS-222</td>
<td>0.61</td>
<td>0.71</td>
<td>0.42</td>
<td>0.61</td>
<td>0.04</td>
<td>0.40</td>
<td>−0.09</td>
</tr>
<tr>
<td>B73</td>
<td>0.41</td>
<td>0.58</td>
<td>0.37</td>
<td>0.56</td>
<td>0.62</td>
<td>0.48</td>
<td>0.07</td>
</tr>
<tr>
<td>Mo17</td>
<td>0.64</td>
<td>0.68</td>
<td>0.77</td>
<td>0.78</td>
<td>0.35</td>
<td>0.56</td>
<td>−0.04</td>
</tr>
<tr>
<td>Average</td>
<td>0.58</td>
<td>0.66</td>
<td>0.62</td>
<td>0.54</td>
<td>0.72</td>
<td>0.34</td>
<td>0.04</td>
</tr>
</tbody>
</table>

ᵃCorrelations between S₇ lines per se and their respective S₈ testcrosses; ᵇCorrelations on diagonal are for S₁ vs. S₈ testcrosses; correlations above diagonal are among S₁ testcrosses; and correlations below diagonal are among S₈ testcrosses
There was no correlation between yields of S7 lines per se and S8 testcross yields (last column, Table 8.11). Genetic correlations between S1 and S8 testcrosses ranged from 0.22 for BSSS and B73 testcrosses to 0.91 for BS13(S)C1 and Mo17 testcrosses at the S1 level and from 0.37 for BSSS-222 and B73 testcrosses to 0.78 for B73 and Mo17 testcrosses at the S8 level. At both levels of inbreeding, the lowest correlation was between a low performance and a good performance tester and the highest correlation was between two good performance testers. Averaged over the four correlations for each tester, BS13(S)C1 at the S1 level and Mo17 at the S8 level had the highest correlations. The individual and average correlations were consistently greater at the S8 level. The correlations for each tester were too small to accurately predict S8 testcross yields on the basis of S1 testcross yields. Correlations for the narrow genetic base testers were greater than for the two broad genetic base testers. Although correlations between S1 and S8 testcrosses were too small for predictive purposes, several S1 testcrosses were consistently ranked high by each tester and S1 testcrosses also were good indicators of their S8 testcrosses (Hallauer and Lopez-Perez, 1979). The S1 testcrosses for BSSS-222 (the poor performance line tester) correctly predicted S8 testcrosses in 34 of 50 instances. The graphic relations of the two sets of testcrosses BSSS-222 (Fig. 8.2) show only one major exception.

The exception was that line 46 in testcross with BSSS-222 yielded 59.1 q/ha as an S1 line and 74.8 q/ha as an S8 testcross, the greatest yielding S8 testcross. The odds, however, were in favor of S1 testcrosses for determining which lines to retain for further selection and testing. Relations were not as good for the other testers, with B73 S1 testcrosses having the poorest and Mo17 S1 testcrosses similar to those for BSSS-222. Because additive effects seemed more important for differences among testcrosses for the five testers, average performance of the lines with the five testers

![Fig. 8.2 Distributions of BSSS-222 (poor performance line tester) S1 and S8 testcrosses](image)

(\(\bar{x}=60.2\) q/ha)
would provide the best measure of combining ability. Figure 8.3 shows graphically the relative $S_1$ and $S_8$ testcrosses averaged for the five testers.

The highest yielding $S_1$ testcrosses generally were the higher yielding $S_8$ testcrosses and seem to support the proposed merits of early testing. Greater correlations between testcrosses of earlier and later generations of inbreeding were reported by Jensen et al. (1983) and Lile and Hallauer (1994) for lines that had undergone selection. Visual selection was done among and within progenies for plant and ear type, flowering date, and root and stalk, but grain yield data of lines per se during inbreeding were not taken. Jensen et al. (1983) for $S_2$ vs. $S_5$ testcrosses ($r = 0.67$) and Lile and Hallauer (1994) for two sets of $S_2$ vs. $S_7$ testcrosses ($r = 0.97$ and 0.86) reported greater correlations than with the use of unselected lines (Hallauer and Lopez-Perez, 1979).

Although some disagreement exists for the appropriate stage of testing for combining ability, it seems that some form of early testing is included in most breeding programs. Early-generation testcrossing may be called early testing but may be delayed until the $S_2$ generation of inbreeding. Although early testing did not imply a perfect relation between the initial and later generations of inbreeding, it was designed to separate the population of lines into good and poor groups for combining ability. Early testing seems to be effective in assigning lines into good and poor groups. As a consequence, greater emphasis for selection and testing can be placed on better lines. Yield testing, however, is expensive in money, time, and energy at whatever stage of inbreeding it is used for making the initial testcrosses. Early testing was designed to reduce effort on progenies that have poor combining ability, but sampling may be limited in the number of crosses that can be produced and tested. Alternatively, a larger sample of $S_1$ progenies of selected $S_0$ plants can be
produced and screened visually for resistance to pests and for morphological traits unacceptable in lines for use in producing hybrid seed. Selection among and within \(S_1\) progenies and advancing to the \(S_2\) generation by self-pollination can be done relatively easily. In a way, early testing might be similar to testing lines developed through doubled-haploid technology.

A compromise between early- and late-generation testing would be possible for lower cost breeding programs. Additional selection among and within \(S_2\) progenies can be continued and initial testcrosses made in the \(S_3\) generation. If, for example, 500 – 600 \(S_1\) progenies can be produced from a breeding population, selection during \(S_1\) and \(S_2\) generations may reduce the number for testcrossing to, say, 100 in the \(S_3\) generation. In contrast, it may not be possible to produce and test 500 – 600 testcrosses of selected \(S_0\) plants. Although correlations between line and hybrid traits may be low and effectiveness of visual selection for combining ability questioned, combining ability tests at the \(S_3\) generation would be for a group of phenotypically elite lines that can be propagated as lines and useful as parental stocks in hybrids. Lack of a genetic correlation between selected traits of the lines and their combining ability would be an advantage rather than a disadvantage for many traits. Thus one would expect an approximately normal distribution for combining ability of phenotypically desirable lines, provided the lines are from the same population. Other traits, stalk quality for example, are very important in all types of circumstances in which maize is grown. Jenkins (1929) and Sprague (1946) found correlations of 0.88 and 0.98 for stalk quality between \(S_0\) and \(S_1\) generations, which is excellent, on the assumption that a negative correlation does not exist between stalk quality and combining ability for yield. Eberhart (1974), however, failed to detect any significant correlations between stalk quality and combining ability.

The final analysis of testing determines the effectiveness of testing methods used in identifying lines that have good performance and extensive use in hybrids. Of the 27 lines studied by Jenkins (1934), 2 lines (L289 and L317) of the Lancaster series were used extensively in hybrids. L289 and L317 ranked in the upper half of the lines tested and would have been selected with early testing. Two other lines (I224 and L304A) included in Jenkins’ study were used to a limited extent. I224 had the greatest yields in the Iodent \(S_1\) series and L304A was in the upper 50% of the Lancaster \(S_1\) series. From the data reported by Sprague (1952), three lines were identified by early testing that had some use in commercial hybrids; B10 and B11 had limited use, whereas B14 was used extensively. A fourth line B37 also was later released and used extensively. B73 and B84 were released for use in commercial hybrids. Both lines were initially selected on the basis of early testing in recurrent selection programs. Both B73 and B84 exhibited high yield in crosses with Mo17. Therefore, evidence indicates early testing is capable of identifying lines that have eventually had use in hybrids. Of course errors have been made in selecting lines on the basis of early testing, but the early discarding of lines permits savings in time, money, and effort expended on lines tested at more advanced stages of inbreeding as opposed to doubled haploids. Greater effort in selection and testing can be directed to those lines that have survived early testing.
Bernardo (1991) determined theoretically that the efficacy of early-generation testing was limited primarily by non-genetic effects. He found that the genetic correlation \( r_{GnGn'} \) between testcross performance of \( S_n \) and \( S_{n'} (n' > n) \) individuals or lines was the ratio of their testcross genetic variances, which was a function of the inbreeding coefficients \( (F) \) of the two selfing generations of the lines being testcrossed:

\[
r_{GnGn'} = \left[ \frac{(1 + F_n)}{(1 + F_{n'})} \right]^{0.5}.
\]

Based on this relation, Bernardo (1991) determined the genetic correlations between lines at generation \( n \) with their directly descended lines at generation \( n' \) (Table 8.12).

### Table 8.12

Expected genetic correlations \( (r_{GnGn'}) \) between testcrosses of \( S_n \) and \( S_{n'} \) maize lines and correlations between testcross phenotypic value at generation \( n \) and true genetic value \( (r_{PnGx}) \) at homozygosity. The \( r_{PnGx} \) values are for three heritability estimates of \( S_0 \) testcrosses.

| \( S_n \) line | Inbreeding coefficient \( F_n \) | \( S_{n'} \) line | \( r_{GnGn'} \) | \( r_{PnGx} \) | \( \hat{h}^2 \) 
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>0.0</td>
<td>( S_2 )</td>
<td>0.82</td>
<td>0.35</td>
<td>0.25</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>0.5</td>
<td>( S_3 )</td>
<td>0.76</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>0.75</td>
<td>( S_4 )</td>
<td>0.73</td>
<td>0.57</td>
<td>0.75</td>
</tr>
<tr>
<td>( S_4 )</td>
<td>0.875</td>
<td>( S_5 )</td>
<td>0.72</td>
<td>0.60</td>
<td>0.86</td>
</tr>
<tr>
<td>( S_5 )</td>
<td>0.9375</td>
<td>( S_6 )</td>
<td>0.71</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>( S_6 )</td>
<td>0.96875</td>
<td>( S_x )</td>
<td>0.71</td>
<td>0.62</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Adapted from Bernardo (1992)

The heritability estimates of the testcrosses are important for determining the relative correlations of the different generations for retention of the genetically superior lines (Bernardo, 1992). Heritability estimates can be enhanced with better experimental techniques, experimental design, and extent of testing. Heritability estimates of testcrosses (half-sib families) evaluated in two replications at three to four environments are usually in the range of 0.45 – 0.60 for \( S_0 \) individuals. Lower heritability estimates have a greater impact on the correlations between testcross phenotypic value at generation \( n \) and the true genetic value at homozygosity (Table 8.12). Based on Bernardo’s relation, genetic correlations range from 0.71 for \( S_1 \) and \( S_x \) to 0.99 for \( S_6 \) and \( S_x \). The genetic correlations, as expected, of adjacent levels of inbreeding are greater than the more distant levels (e.g., \( r_{S_1S_2} = 0.82 \) vs. \( r_{S_5S_6} = 0.71 \)), but the genetic correlations for all combinations suggest early testing should be effective. Bauman (1981) reported that 50% of the maize breeders made the initial testcrosses by the \( S_3 \) generation in the early 1980s. The heritabilities at the \( S_2 \) and
S₃ generation testcrosses would be greater than for S₀ individuals, and testing of fewer lines at the S₂ and S₃ inbreeding levels should correctly identify those lines with above average combining ability. Testing of new lines at earlier generations of inbreeding is much greater than reported by Bauman (1981). Rapid advancements in plot equipment to plant and harvest experimental plots and computer hardware and software to record and analyze data have reduced some of the concerns for testing large numbers of new lines of earlier generations of inbreeding, especially at the S₀ generation.

Development and use of different methods of recurrent selection have emphasized early testing for discriminating among progenies to determine which ones to recombine and form the next cycle of selection. The objective of early testing in recurrent selection is the same as originally proposed for early testing: to identify genotypes that have good and poor combining ability in crosses. Errors will be made in those cases intermediate in combining ability. Evidence given in Chapter 7, however, indicates that long-term cyclical selection programs have been effective for yield improvement.

Early testing, in the purest sense, was used in nearly all instances. Although Payne and Hayes (1949) and Hayes (1963) questioned the value of early testing, Hayes (1963) considered recurrent selection important for the future progress of maize breeding and Hayes and Garber (1919) recommended recurrent selection for use in maize breeding. A ceiling on the assemblage of genes is imposed by the particular S₀ plant selected for self-pollination. Recombination of genes permits some additional selection in later generations, but it is minor compared to original selection of S₀ plants. Thus it seems that some form of early testing is desirable to enhance the efficiency of any breeding program. Betran et al. (2004) and Hallauer and Carena (2009) discussed most of the issues related to the importance of testcrosses to determine the genetic worth of newly developed inbred lines as potential parents of elite hybrids.

### 8.6 General vs. Specific Combining Ability

Combining ability of inbred lines is the ultimate factor determining future usefulness and commercial potential of the lines for hybrids. Combining ability initially was a general concept considered collectively for classifying an inbred line relative to its cross performance. Sprague and Tatum (1942) refined the concept of combining ability, and the two expressions of general (GCA) and specific (SCA) combining ability have had a significant impact on inbred line evaluation and population improvement in maize breeding.

Sprague and Tatum (1942) interpreted their results relative to the type of gene action operative for two groups of inbred lines:

1. Six tests of single-cross hybrids involving inbred lines that had survived previous selection and testing.
They analyzed the diallel crosses to determine the relative importance of GCA and SCA for the lines included in each set of crosses. Although the diallel had been used before in maize breeding to determine the yield of inbred lines in crosses, Sprague and Tatum (1942) apparently were the first to partition the total combining ability of the lines into GCA and SCA. They defined GCA as the average performance of a line in hybrid combinations and SCA as those instances in which certain hybrid combinations are either better or poorer than would be expected on the average performance of the parent inbred lines included. They also emphasized that estimates of GCA and SCA are relative to and dependent on the particular set of inbred lines included in the hybrids under test, an important principle that is often forgotten. Sprague and Tatum (1942) found that GCA was relatively more important than SCA for unselected inbred lines, whereas SCA was more important than GCA for previously selected lines for influencing yield and stalk lodging. Also, they interpreted GCA as an indication of genes having largely additive effects and SCA as indicative of genes having dominance and epistatic effects. Their results supported usage of the testcross test for preliminary evaluation of inbred lines for GCA, although single-cross trials are necessary to determine the most productive specific combinations. Although GCA can be determined from single-cross combinations, it can be more effectively determined with testcross tests, particularly if preliminary information is needed for a large number of lines. The conclusions of Sprague and Tatum (1942) are equally valid in present breeding programs because of interest in production and growing of single crosses.

The notions of GCA and SCA have been used extensively in breeding of maize and other crop species. Types of genetic materials and methods used, traits studied, and interpretations made have been diverse. Some problems associated with diallel analysis of crosses for estimating GCA and SCA were emphasized in Chapter 4. Statistical geneticists have made important contributions to clarifying genetic information from an analysis of a diallel set of crosses. Henderson (1952) and Griffing (1953) defined and applied GCA and SCA concepts to animal and plant experiments. Hull (1946, 1952), Griffing (1950), Hayman (1954a, b), and Jinks (1954) gave procedures for estimating the genetic parameters for restricted genetic models. Griffing (1956a) showed the relation of the diallel cross-mating design to the concept of covariances between relatives in terms of additive and dominance genetic effects. Griffing (1956b) presented the analysis of variance for four situations of the diallel mating design. Matzinger et al. (1959) showed how GCA and SCA and their interactions with environments can be calculated and interpreted for the diallel mating design. In most if not all instances the estimates from the diallel mating cannot be interpreted relative to a reference population (see Chapter 4).

The concepts of GCA and SCA became useful for characterizing inbred lines in crosses and often were two of the traits included in the description of an inbred line. Later developments in the characterization of genetic variance and types of gene action operative in crosses of inbred lines also often were interpreted relative to
GCA and SCA of inbred lines. Different selection methods proposed for recurrent selection also were considered in the context of GCA and SCA and the type of gene action contributing to the heterosis expressed in crosses. If additive gene effects with partial to complete dominance are important, recurrent selection methods that emphasize GCA should be used (Jenkins, 1940). If overdominance is of primary importance, recurrent selection methods that emphasize SCA would be appropriate (Hull, 1945). Comstock et al. (1949) designed the RRS method to enhance selection for all types of gene action (i.e., both GCA and SCA). Before development of the RRS method, maize breeders often were polarized relative to the importance of GCA and SCA.

Experimental results listed in Chapter 7 indicate that selection based on either broad base or narrow base testers generally has been effective for the improvement of both the population and the testcross performance. It was commonly accepted that the use of a narrow base tester would improve the combining ability to the specific tester but would have little value for the improvement of GCA, which is based mainly on additive genetic effects. However, research has shown that the use of inbred lines as testers has improved both combining ability for specific testers and GCA as measured by the performance of populations themselves or of testcrosses with unrelated broad base populations (Horner et al., 1973; Russell et al., 1973; Russell and Eberhart, 1975; Hoegemeyer and Hallauer, 1976; Horner et al., 1976; Walejko and Russell, 1977). The results seem logical in view of the summaries for relative types of genetic variance (see Chapter 5) and Sprague and Tatum’s (1942) original conclusion that GCA is more important for previously unselected material.

In selection experiments, the testcross progenies (usually half-sibs) are the original cross information used to select those progenies for recombination. Horner (1963, 1976) reported a study testing Hull’s hypothesis of the effectiveness of recurrent selection for SCA and speculated that about one-third of the gain from use of an inbred tester was specific to the tester and about two-thirds was due to the additive effects transferable to other combinations. Horner (1976) also concluded that maize breeders should be able to change testers as the occasion demands because overdominance and epistasis do not seem to be of major importance in yield heterosis. Horner’s (1963, 1976) conclusions were supported by Russell et al. (1973) and Walejko and Russell (1977) for different populations and inbred testers.

The conclusions for the selection studies were corroborated by Russell and Eberhart (1975) and Hoegemeyer and Hallauer (1976). The former compared selected lines from two recurrent selection programs that evaluated half-sib progenies. Analysis of variance of the crosses among lines indicated that most of the variation within each set of crosses was attributable to the average performance (or GCA) of the lines included in the crosses. In five of six instances, the GCA mean square was highly significant and relatively large compared with the SCA mean square, indicating additive gene effects were the primary type of gene action operative in the crosses. Hoegemeyer and Hallauer (1976) tested single crosses among S7 lines selected by a full-sib selection scheme that emphasized selection for SCA. Crosses among the elite lines that had survived selection for four generations of selection for SCA also had high GCA with other elite lines. Hoegemeyer and
Hallauer (1976) also showed that their selection procedure was effective for SCA, but the SCA effects must be small relative to the GCA effects. Although present evidence seems to show that GCA (or additive gene effects) is more important than SCA, the terms still are commonly used in the sense of the maize breeding jargon and will continue to be so. Maize breeders breed and test to identify that unique combination of inbred lines for high productivity. SCA is expressed although the GCA was probably more important in identifying the lines for the unique combination. Non-additive gene effects seem to be small on average, but they may be important for that one unique combination. Heterotic effects are unique for each hybrid. Therefore, breeding methodologies and genomic tools targeted at conductive effective selection should be aware of this uniqueness. For instance, sequencing efforts on only B73 may limit the identification of useful alleles for combining ability and other complex traits.

References

Eckhardt, R. C., and A. A. Bryan. 1940a. Effect of method of combining the four inbred lines of a double cross of maize upon the yield and variability of the resulting double crosses. J. Am. Soc. Agron. 32:347–53.
Eckhardt, R. C., and A. A. Bryan 1940b. Effect of the method of combining two early and two late inbred lines of corn upon the yield and variability of the resulting double crosses. J. Am. Soc. Agron. 32:645–56.


Hayes, H. K. 1963. *A Professor’s Story of Hybrid Corn.* Burgess, Minneapolis, MN.


References


Chapter 9
Inbreeding

9.1 The Need for Maize Artificial Pollination

Maize is a naturally cross-pollinated crop and dispersion of pollen grains (male gametes) is achieved by wind currents, system that favors cross-pollination. The plant has a separate male and female inflorescence (monoecious) which makes relatively easy to produce seed by artificial hybridization and self-pollination for inbreeding. The tassel (male inflorescence with staminate flowers) is at the top of the plant arising from the shoot apical meristem. The ear (female inflorescence with pistillate flowers) is usually located in the middle of the stalk (at the sixth of seventh node from the top of the plant) and originates from the axillary bud apice. During development flowers become unisexual. Apical dominance is present on stalks with multiple ears (prolific genotypes). The male florets usually mature before than the female florets (protandry). Genotype and environmental conditions (e.g., stress) influence the difference in maturity of male and female florets. Pollen shed occurs after anther exertion from each spikelet and begins in the main tassel branch (central spike or rachis). Each spikelet has two florets and pollen shed starts from the upper flower. Spikelets are in pairs: pedicellate and sessile. Each spikelet has a pair of glumes. Within the glumes each floret is also enclosed with thin scales (a lemma and a palea). Two of the three anthers are located adjacent to the palea and the third one is located adjacent to the lemma. The number of pollen grains dispersed by the tassel depends on the genotype and/or vigor of the plants. Hybrids, for example, shed more pollen than inbred-lines. Some open-pollinated varieties, however, shed even more pollen than hybrids. Ear shoots are formed of husks (modified leaves) and silks emerging from the cob. The silks are functional stigmas and there is one stigma for each potential kernel. Silk emergence progresses from the bottom to the tip of the ear. High temperatures and low-moisture availability may cause silk growth to stop and not be ready for fertilization at the time of pollen shed. The ear branch or shank is formed of nodes and short internodes. The ear has a thick axis (cob) which is similar to the tassel rachis since it produces multiple rows of paired spikelets. Each spikelet is enclosed by a pair of glumes and contains only one functional floret. Two florets are initiated but only the upper one is developed. Both florets produce a lemma but only the upper one produces a palea. Each upper floret produces an ovary with an elongated style (silk) covered with trichomes (hairs).
Fertilization is accomplished when pollen grains come in contact with the silk, grow down the styles, and unite with the female gametes. Double fertilization in the embryo sac results in the formation of the embryo \((2n = 20)\) and the endosperm \((3n = 30)\) of the seed. Poor seed set is the consequence of non-viable gametes or poor synchronization of the time of pollen and silk emergence.

Artificial pollination, whether hybridization or self-pollination, allows the breeder to maintain the parentage of the seed produced. Techniques have been developed in which inexperienced individuals can self and cross selected plants and produce 300–500 seeds per pollination (Russell and Hallauer, 1980). Both the male and the female inflorescence needs to be covered with the purpose to prevent contamination (tassel and shoot bags). There is a sequence of methods used in making self and cross-pollination that should be followed to produce good-quality and contamination-free seed. The first step is the use of shoot bags to cover the female gametes. Ear shoots must be covered before the silks have emerged to avoid fertilization by pollen of unknown origin (e.g., if even a few silks have been exposed do not cover). This is the most important phase of artificial pollination and plants should be checked daily at flowering time. The second phase is to examine the tassels and determine if viable pollen is present. The tassels that are used for pollination can be on the same plant as the ear shoot covered (self-pollination) or on different plants (cross-pollination). After identifying that tassels and ears are at adequate maturity they are prepared for pollination (third step). The preparation of tassels and ears, however, is always performed on the day before pollination. The ear tips are cut and tassels bags are used for the source of pollen. Pollination is carried out one day after male and female flowers have been prepared and this step is performed when pollen is available in the field. For a successful pollination tassel bags should be dry and need to be tapped before it is removed from the plant. This step should be done as quickly as possible so that the chances of contamination are minimized. The tassel bag is secured over the ear shoot and fixed to the stalk until harvest. The pollination date is usually identified on the bag. Additional identification is needed for identifying plants to be used (e.g., full-sib recurrent selection).

Because maize is a cross-pollinated species maize populations or varieties include a mixture of genotypes, each unique, that depends on the particular combination of the two gametes that united to form the zygote that produced this particular genotype. The structures that contain the male (tassel) and female flowers (ears) are prominent and facilitate hand pollination. Although the tassel and ear of maize are distinctive and easily recognizable from other crop species, variations in the tassel and ear occur in races, populations, varieties, hybrids, and inbred-lines. Tassels may differ in length, number and spacing of tassel branches, compactness, color, ease of pollen shed, amount of pollen shed, etc.; all characteristics required when applying for plant variety protection (PVP) and intellectual property. Ears vary in relative placement on the plant, length, diameter, kernel-row number, ear number, kernel color, etc.

The protandrous nature of maize promotes cross-pollination, but a small amount of self-fertilization may occur when there is some overlap of pollen shed with silk receptivity. In a few examples protogyny occurs (Weatherwax, 1955). Kiesselbach
(1922) reported only 0.7% self-pollination occurred in Nebraska White Prize plants grown in a field of Hogue’s Yellow Dent. He concluded from his data that the amount of self-fertilization that occurs under ordinary field conditions for maize is negligible.

Departures from randomness of mating have been reported. They occur because of differences in maturity of individual plants in a population and environmental conditions that exist at flowering time. Gutierrez and Sprague (1959) reported that maize exhibited marked departures from randomness of mating. They also concluded that dates of silking and pollen shedding, number of plants shedding pollen, length of pollen shedding period, plant height, and possibly selective fertilization (Jones, 1924) may have been factors contributing to departures from randomness of mating. Some of these factors would tend to develop subpopulations within a presumably random mating population. These subpopulations are formed because, for instance, early flowers would tend to mate with early flowers and late flowers would tend to mate with late flowers. Although some physiologically isolated subpopulations would exist within an open-pollinated maize population, there also would tend to be cross-pollination among subpopulations, sufficient to prevent easily distinguishable subpopulations.

Although complete randomness of mating may not be realized in an open-pollinated population of maize, the genotypic array within a maize population is extensive and complex. As stated before each individual within a population arises from the union of two different gametes. Hence each plant is in reality a different F₁ hybrid. An array of all the plants (hypothetical F₁ individuals) in a population would approximate a normal distribution for quantitatively inherited traits. For instance, there would be a range in distribution for each trait but the population would be characterized by the plants within two standard deviations of the mean. A maize population, therefore, would include individual plants that are as productive as hybrids produced from crosses of inbred-lines. The problem is one of identification and perpetuation. Not until studies were conducted to reduce the complex genotypes of individual plants of maize populations to pure genotypes (inbred-lines) that were reproducible and useable as parental stocks for the production of hybrids was the modern concept of maize breeding developed.

Inbreeding occurs, either naturally or artificially, by mating individuals that are more closely related than by random chance. These individuals would be closer than the average relationship within the particular population defined as non-inbred. We are concerned with effects of inbreeding that occur from controlled matings, but inbreeding can occur naturally in small populations because of the limited number of individuals and, therefore, the limited number of matings that can occur. Occurrence of inbreeding as a result of population size rather than by controlled matings is of particular concern in maize breeding programs that include recombination of selected individuals in various recurrent selection procedures, because a limited number of individuals is selected for recombination and random drift in allele frequencies could occur if effective population sizes are low.

Inbreeding is determined by the percentage of homozygosis of the zygote as inbreeding is a description of the zygotes and not the gametes. Gametes from a
homozygous individual, such as an inbred-line of maize, are no more or less inbred than gametes from a completely heterozygous individual, such as an F₁ hybrid, between two relatively homozygous inbred-lines. This relation indicates why individual plants in a cross-pollinated population of maize can be considered as unique F₁ hybrids as they each were formed from the union of two gametes that are not identifiable and repeatable. Hence inbreeding, primarily by self-fertilization, is important in the modern concept of maize breeding.

9.2 Early Reports of Inbreeding

The early history of inbreeding and hybridization was reviewed and summarized by East and Jones (1918) and Jones (1918). Kölreuter (1776), Knight (1799), Gärtner (1849), and Focke (1881) conducted extensive experiments with many plant species and were prominent hybridizers before the rediscovery of Mendel’s laws of segregation and recombination. None of them, however, realized that inbreeding and outbreeding were opposite expressions of the same phenomenon. Darwin (1877) also recorded many observations of the effects of inbreeding and outbreeding in plant species, but his interpretations also were made before Mendelism was understood. Darwin thought that inbreeding was not a normal process of mating and that the effects of inbreeding accumulated until eventually the species was doomed to extinction. However, he made one important interpretation of inbreeding effects. These effects usually had been attributed to the inbreeding process itself rather than to homozygosity. Darwin observed that it made no difference in vigor whether plants in an inbred lot were selfed or crossed among them. He attributed this to the members of an inbred lot becoming germinally alike. This interpretation was supported by crosses of his selfed lines with other lines inbred to a less degree that did not express as great an increase in vigor as the crosses of the same lines with a fresh stock from different regions. Crosses between two inbred-lines, however, did give a noticeable increase in vigor, often exceeding the original variety. Differences in the performance of inbred-lines and their crosses were attributed to germinal similarity and diversity. Darwin’s interpretations were similar to our concept of segregation and recombination, and many of the problems that bothered Darwin would have been clarified if knowledge of Mendelism had been available to him.

Although the phenomenon of inbreeding and crossing was studied extensively in the plant kingdom, not until the studies reported by Shull (1908, 1909, 1910) in maize was the fundamental similarity of inbreeding and crossing clearly understood. Moreover, not until Shull’s (1908) interpretations became available was the exact genetic basis of inbreeding understood. East (1908, 1909) observed similar phenomena. But East (1909) also was apprehensive about the practical utility of inbreeding to develop parental stocks to produce crosses for the improvement of maize production.

Although Darwin had conducted some inbreeding studies with maize, the beginnings of the modern concept of maize breeding methods (inbred-lines and hybrids)
are accredited to Shull and East. Long-term inbreeding studies in maize were few and most of the results were not encouraging. Shamel (1905), Collins (1909), and Davenport and Holden (cited by Shull, 1952) reported on the effects of inbreeding in maize. In most instances the negative effects of inbreeding were recognized but the advantages of its use in crosses were not. The studies of G. H. Shull and E. M. East that discovered and revealed the potential of hybrid corn were conducted independently at the same time. A historical perspective of the studies by Shull and East was given by Shull (1952) and Hayes (1956).

The inbred–hybrid concept that made hybrid maize a reality was defined and applied first in the public sector and traces back to the early 1900s. East, later replaced by H. K. Hayes and D. F. Jones, led one of the research groups that discovered and revealed the potential of hybrid maize using Leaming inbred-lines. East was directly influenced by the biological principles of Darwin, Mendel, and Vilmorin in relation to plant improvement (Hayes, 1956). East related those principles to the more practical plant improvement studies to achieve his goals. But probably the most well-known public scientist in early hybrid maize research is Shull whose maize research records date back to 1904. At that time research focused on heredity as a basis for improving plants and animals and Shull studied the theory of genetics and its application to plant breeding. East at Connecticut and Shull at Cold Spring Harbor independently started studies of inbreeding and crossbreeding in 1906 (Hayes, 1963) and provided essential insight into the efforts of maize inbreeding. However, Shull discontinued his studies in 1916 since he concluded the inbred–hybrid concept had no practical value due to small amount of seed produced on inbred-lines. Jones suggested to East a procedure that would make hybrid corn a reality for industry and farmers, using the already developed Leaming lines as females and Burr’s White inbreds as males (Jones, 1918). The change from open-pollinated varieties to double-cross hybrids was a significant improvement in developing maize hybrids with improved standability and grain yield performance. It had, however, a negative impact on breeding efforts toward population improvement and genetic diversity (Carena, 2007).

The correct interpretation of the phenomenon of inbreeding and crossing was given by Shull while the practical implications of the inbred–hybrid concept were emphasized by East. Shull’s remarkable conclusions were based on a limited amount of information, but they formed the basic principles of maize breeding as we know them today. Shull’s studies in maize were designed not to study the effects of inbreeding and crossing but rather the effects of the phenomenon on the inheritance of number of kernel rows per ear. Based on the inheritance of kernel rows per ear, plant height, and yield of the inbred-lines developed by selfing and their performance in crosses among the inbred-lines, Shull (1908, p. 299) concluded: ‘The obvious conclusion to be reached is that an ordinary cornfield is a series of very complex hybrids produced by the combination of numerous elementary species. Self-fertilization soon eliminates the hybrid elements and reduces the strain to its elementary components. In the comparison between a self-fertilized strain and a cross-fertilized strain of the same origin, we are not dealing then, with the effects of cross- and self-fertilization as such, but with the relative vigor of biotypes and
Shull’s conclusion that self-fertilization separated hybrids into pure forms also was an important contribution to what is now termed pedigree selection. Specific crosses are made between inbred-lines and self-fertilization is imposed on the F1 hybrid in an attempt to derive inbred-lines that are superior to either parent because of genetic segregation and recombination (e.g., transgressive segregants). Shull (1908) emphasized that ‘there is no intelligent attempt in these methods to determine the relative value of the several biotypes in hybrid combination, but only in the pure state. In the present state of our knowledge it is impossible to predict from a study of two pure strains what will be the relative vigor of their hybrid offspring.’ The insight of these conclusions is interesting because maize breeders are still faced with determining the relative merits of inbred-lines in hybrids (see Chapter 8).

Shull (1909) discussed further the results of his studies and outlined the two essential features for the production of hybrid maize. He referred to the process as the pure line method in maize breeding, which he considered as having two stages:

1. Finding the best pure lines, and
2. The practical use of the pure lines in the production of seed corn.

He discussed the necessity of making as many self-fertilizations as possible and continuing self-fertilization until the lines were in the homozygous state. He correctly hypothesized that lines developed by self-fertilization would not be useful for all situations and that the only method for determining relative merit of the lines was to make all possible crosses among the lines and evaluate them in F1 hybrids. He concluded that after one has determined which pair of pure strains (pure lines) produce the desired results, the method of producing the seed is relatively simple but it may be somewhat costly.

Shull (1910) reinforced his arguments for the advantages of the pure line method of maize breeding. The salient feature of his discussion was the comparison of the pure line method with varietal hybridization methods discussed by East (1909) and Collins (1909). Shull reasoned correctly that only by reducing the complex mixtures of genotypes in a variety to their pure forms and determining the best pair in F1 hybrids would maximum uniformity and performance be attained. Because varieties were complex mixtures of genotypes, the modal level of heterozygosity of the variety cross would be less than the cross of two pure lines. Recurrent selection methods, however, have been able to improve the frequencies of favorable alleles in populations at a level they do not only provide elite inbred but also offer population hybrids that can be profitable alternative seed production systems for areas where investment for inbred-line development programs and single-cross hybrid production systems is
9.3 Inbreeding Systems

Inbreeding is a result of mating individuals of some degree of resemblance. Extreme and non-extreme forms of inbreeding play important roles in plant breeding. For instance, inbreeding increases the genetic variability among individuals in a population allowing more efficient genetic improvement. Also, a slower fixation of undesirable alleles allows more opportunities for selection during inbreeding. Therefore, breeding programs must weight the importance of selection against the length of time required to achieve a level of inbreeding. However, selection for desirable alleles reduces the potential negative effects inbreeding can cause.

Inbreeding in maize usually is by self-fertilization because of the rapid approach to homozygosity. Because maize has the male and female organs in separate inflorescences, it is easy to manipulate self- and cross-fertilizations artificially. Self-fertilization is the most extreme form of inbreeding, but other methods are available. Other less restrictive forms of inbreeding have been suggested for producing more vigorous inbred-lines (Macaulay, 1928; Lindstrom, 1939; Harvey and Rigney, 1947; Kinman, 1952; Stringfield, 1974). Theoretically, less restrictive forms would permit less rapid fixation of deleterious genes than the selfing method. The zygotic array of the progeny by selfing is determined by the gametic array of the plant that is selfed. Future selection during inbreeding would be fixed within the limits of the genotype of the original $S_0$ plant selfed. If, for example, some form of sib-mating is used to effect inbreeding, fixation of deleterious alleles would be slower and opportunities for selection would be enhanced. The advantages and disadvantages of intense inbreeding (i.e., selfing vs. a milder form) have to be considered. For instance, do the advantages of selection during inbreeding outweigh the advantages of rapid approach to homozygosity by selfing?

Some of the common systems of inbreeding are listed in Table 9.1. Earlier methods for measuring expected average level of inbreeding were suggested, but
Table 9.1 Coefficients of inbreeding (expected homozygosity) in different generations for different systems of inbreeding

<table>
<thead>
<tr>
<th>System of inbreeding</th>
<th>Backcrossing$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F of parent = 0</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
</tr>
</tbody>
</table>

$^a$ Also the same as crossing the offspring to the younger parent

$^b$ Involves the crossing of one parent to its offspring

$^c$ Recurrent relation for calculating expected homozygosity in the system of inbreeding where $F$ is Wright’s coefficient of inbreeding, $F'$ refers to the previous generation, and $F''$ refers to the second generation removed

formulas for measuring inbreeding from matings between different kinds of relatives were first derived by Wright (1921, 1922a). Wright’s coefficient of inbreeding ($F$) is the level of correlation between uniting gametes and measures directly the probable increase in homozygosis. Although Wright (1922a, b) was the first to derive such formulas, others have obtained the same results with different methods. The method derived by Malécot (1948) has been the one most widely used as it is easier to visualize and gives the same results as the path coefficient method of Wright (1922a). Malécot defines the coefficient of inbreeding as the probability that two genes at the same locus are identical by descent. Two allelic genes at a locus may be the same because they are either alike in state or identical by descent. Alleles that are alike in state have the same function, but so far as known they do not occur together because of ancestry unless they originated from some unknown remote ancestor. Hence, the coefficient of inbreeding, or level of inbreeding, of relatives is relative to a specified base population.

Theory, use, and illustrations for determining the coefficient of inbreeding of relatives for different systems of inbreeding are given by Kempthorne (1957), Pirchner (1969), and Li (1976). The coefficient of inbreeding, therefore, is computed to determine the level of homozygosity in a specified generation. The approach to homozygosity varies among the different systems of inbreeding. Self-fertilization approaches homozygosity faster than any of the other systems, but in plant species that are dioecious or have self-incompatibility systems (and in higher animal species having separate sexes) it is not possible to self-fertilize; hence, full-sib mating is the closest form of inbreeding possible. Maize, however, is amenable to all systems of inbreeding, and the system used depends on either the objectives of the breeding program or the fundamental studies under investigation.
The advantages of self-fertilization to develop inbred-lines that are relatively homozygous are obvious in Table 9.1. It requires four generations of full-sib matings and six generations of half-sib matings to have the same theoretical level of inbreeding, or homozygosity, as one generation of self-fertilization. Even if the off season (e.g., winter nurseries for the US Corn Belt) is available and used, the number of seasons required to attain the same theoretical level of homozygosity will be greater. Half-sib and full-sib systems of inbreeding also will usually require one additional season to develop the families for continued matings. If one wants to use the full-sib method of inbreeding in a particular population, it is necessary to make controlled pollinations to develop full-sib families. Seed harvested from plant-to-plant crosses (full-sibs) will be non-inbred and cross-pollination of the full-sib individuals in the next season will yield seed that has an expected homozygosity of 25%; in the same two seasons, the expected homozygosity by selfing is 75%.

Half-sib families can be obtained in two ways, depending on the population to be used to initiate inbreeding:

1. Half-sib seed can be obtained from an isolation field by harvesting and shelling each ear separately. Assuming no natural self-fertilization, female gametes of each plant will be pollinated by a random sample of male gametes from the population; hence the seeds are half-sibs having a common female parent.
2. In the nursery half-sib progenies are established by controlled pollinations by bulking pollen hoped to be representative of the population and applying the pollen to the silks of a separate sample of plants. In subsequent generations, inbreeding is established by bulk pollination of a plant within the half-sib progeny. In either instance the theoretical level of inbreeding after two generations is 12.5% compared with 75% after two generations of self-fertilization.

Backcrossing usually is not considered one of the common systems of inbreeding, but backcrossing is used extensively either to transfer a specific gene from one genotype to another or to improve an inbred-line for some quantitative trait. If the parent used for backcrossing is homozygous \((F = 1)\) and a simple dominant gene is being transferred by backcrossing \([\text{e.g., } Ht \text{ gene for Exserohilum turcicum (Pass.) Leonard and Suggs.}]\), the rate of inbreeding is the same as selfing (Table 9.1) and the genotype is theoretically rapidly recovered. The theoretical level of inbreeding by backcrossing to a non-inbred parent \((F = 0)\) never exceeds 50%. If complete homozygosity is desired from a backcrossing program that includes a non-inbred parent, some other form of inbreeding is necessary. Homozygosity is eventually achieved by all systems of inbreeding except backcrossing to the non-inbred parent. Improvement of an inbred line for a quantitative trait for say European corn borer resistance or components of yield is less successful because of the dilution effects in successive backcrosses (Geadelmann and Peterson, 1978). However, the use of backcrosses in quantitative traits that are less complex such as flowering time has been extremely successful. The conversion of late-maturing elite lines to early versions through the use of one backcross (Carena et al., 2010) to several backcrosses...
(Rinke and Sentz, 1961; Hallauer et al., 1988, Table 8.3) has expanded maize production to higher latitudes with inclusion of unique and competitive temperate and tropical germplasm.

For maize, we have a choice of inbreeding system, but in many plant species and higher animal species the choice is restricted. Macaulay (1928) suggested an inbreeding system less severe than self-fertilization, which he called plot-inbreeding. He suggested planting seed from selected ears in plots containing 200 to 250 plants, each plot isolated from the others. Within each plot only large uniform ears on selected plants are retained each generation for continued selection. The suggested procedure seems to be a form of half-sib inbreeding. Although no data were presented, claims were made that more vigorous lines were obtained by plot-inbreeding than by self-fertilization and that the best lines developed by self-fertilization must be inferior to the best lines from plot-inbreeding because

1. fifty percent of all heterozygous factors are reduced to homozygosity in one generation by self-fertilization, whereas it takes several generations in plot-inbreeding. Thus there is more chance for selection under plot-inbreeding.
2. lines developed from self-fertilized plants can only inherit favorable factors that are present in the parent plant, whereas plot-inbreeding provides for the gradual accumulation into a single strain all the favorable factors possessed by the individuals in a plot.

Macaulay also believed that it was possible to develop homozygous strains as vigorous as $F_1$ hybrids (which are discussed further in Chapter 10). Availability of isolation would limit the use of this system even if all claims were fully justified. It would require many more generations to develop the lines, and their merit in hybrids would have to be established by conventional procedures.

Lindstrom (1939) and Harvey and Rigney (1947) compared inbred-lines developed by full-sib and selfing systems and combinations of full-sibbing and selfing. In both instances the relative number of lines was limited, and they arrived at opposite conclusions. Lindstrom suggested that mild inbreeding should be used at the beginning to prevent rapid fixation of deleterious traits and provide a broader base for selection under diverse environmental conditions. He further theorized that the same results may be attained by using greater numbers in a self-fertilization program, which is an important point. Harvey and Rigney, on the other hand, suggested that it may be better to inbreed intensively at the beginning to select out deleterious factors than, by mild inbreeding, to retain deleterious factors that segregate out in later generations. Their conclusions were based on the high incidence of barrenness that occurred after full-sibbing for three generations and then selfing for four generations. Similar results were reported by Good (1976).

Kinman (1952) also suggested developing lines for use in hybrids by composite sibbing. On the average, however, testcross hybrids involving sibbed lines were not different in performance or variability from those of $S_3$ lines (three generations of selfing) developed from the same source. Sibbed lines were more variable than the related selfed lines ($S_4$) in maturity, number of tassel branches, plant height,
and yield. Loeffel (1971) tested a series of inbred-lines in hybrids after one to four generations of inbreeding to compare:

(1) selfing with sibbing,
(2) selfing followed by sibbing, and
(3) intensity of inbreeding.

He found no appreciable gains in maturity, yield, or other traits that could be attributed to either the level of inbreeding or the choice between selfing and sibbing methods of inbreeding.

The use of milder forms of inbreeding for developing inbred-lines for use in hybrids was reviewed by Stringfield (1974). He discussed the negative aspects of inbreeding by self-fertilization as follows:

(1) Continuous self-fertilization seems too violent.
(2) Homozygous lines bear many marks of lapses in selection.
(3) Homozygous lines are essentially inflexible.
(4) Homozygous lines have disadvantages in production of seed and may limit the final crop.

All aspects are related to the rapid fixation of genes in the homozygous condition by self-fertilization. Stringfield was of the opinion that better and more useful inbred-lines could be developed from a milder form of inbreeding that permitted effective selection (the breeder’s art) concurrently with inbreeding. He proposed assortative mating by chain crossing making selections and pollination between selected plants. The procedure is a modification of the plot-inbreeding system suggested by Macaulay (1928). Stringfield called it broadline development and developed and tested a few lines. The broadline concept has not been extensively used and it does not seem the method has contributed significantly to line development, especially when current public elite lines produce over 6 t/ha (Hallauer and Carena, 2009).

Suggestions for use of systems of inbreeding in maize other than self-fertilization have been made, but they have not been popular. Often the suggestions are philosophical, emphasizing the rigidity of self-fertilization, particularly if the art of plant breeding is important in line development. Because of rapid fixation of genes by self-fertilization, opportunities for visual selection would be limited within progenies for particular plant types. For example, when \( F = 0.5 \), the total genetic variance assuming only additive effects between means of families will be twice as great as that within families (see Chapter 2). After two generations of self-fertilization, \( F = 0.75 \) and total genetic variance among \( S_2 \) families is six times greater than within \( S_2 \) families. Inbreeding increases effectiveness of selection among families but decreases effectiveness of selection among individuals within a family which is evident by the overall uniformity shown by lines already at the \( S_2 \) stage of inbreeding in nursery fields. Visual selection (Chapter 8) is important for some traits, but visual selection among progenies rather than within progenies is much more effective at that stage. Breeders use visual selection where possible and inbreed to
homozygosity as quickly as possible (e.g., by self-fertilization). Visual selection is improved by duplicating early-generation lines across locations, densities, and planting dates (see Table 1.5, Chapter 1; Carena and Hallauer, 2001; Carena and Wanner, 2003, 2010; Carena et al., 2003; Carena, 2008; Hallauer and Carena, 2009; Sezegen and Carena, 2009). The differences certainly will be greater among progenies than within progenies and recycling of desirable factors can be accomplished by intercrossing desirable genotypes that also have had some testing for combining ability and starting another cycle of inbreeding and selection.

The reasons for use of milder systems of inbreeding have some validity, but the time required to attain a level of homozygosity acceptable to current standards of uniformity for use as parental stocks in hybrids discourages their use. Applied breeders particularly are under pressure to develop new lines for use in hybrids, and they resort to self-fertilization because of its rapid approach to homozygosity. Recycling of new lines and additional pedigree selection is initiated by self-fertilization. Hence the general philosophy is to screen greater numbers of lines developed by self-fertilization than to patiently select fewer lines by inbreeding systems that have a slower fixation of genes. Right or wrong, this system is currently used and will undoubtedly continue to be used in the future. Self-fertilization by pedigree selection certainly contributes to narrowing the genetic base of parental stocks used to produce hybrids. It is not the system of inbreeding but the parental source material that contributes to the problem.

9.4 Inbreeding Due to Small Population Size

The effectiveness of any selection method depends on the balance of the two main forces affecting allele frequencies in breeding populations: selection and genetic drift (Stojsin and Kannenberg, 1994). An increased frequency of favorable alleles is the consequence of selection response. This can be achieved not only by changing additive effects but also with the complement of dominant effects, although effects are not always complementary. Maize breeders conducting recurrent selection programs are continually faced with decisions concerning the number of progenies of a population for evaluation and the number of progenies to select for recombination to form the next cycle of selection for population genetic improvement to be used as improved sources of inbred-lines. The number of individuals sampled for evaluation and recombination is recurring in each cycle. In most instances greater numbers dictate compromises of resource allocation in breeding programs to maintain continuous recycling of material. The populations included are usually closed populations because the researchers are interested in determining the long-term effects of the specific selection programs (see Chapter 7). Eventually all members of a closed population become related to one another depending on the size of the population. Inbreeding is avoided as much as possible in recombining and random mating of the resynthesized and improved populations. But because of the number of individuals that can be reasonably included, inbreeding is inevitable. Once the locus becomes
fixed (or inbred) within a closed population, the only recourse is to outcross to an unrelated population. In small populations, which are common in recurrent selection programs, the closed population cannot be retained without some inbreeding.

Small population sizes cause a random change in allelic frequencies (e.g., genetic drift). In most cases, genetic drift will reduce response of selection although it can act in the same direction of selection in exceptional cases. As a consequence, the frequencies are changed at random resulting in inbreeding depression. Inbreeding depression is the reduction in performance caused by the expression of undesirable alleles. Inbreeding, however, is not the cause of inbreeding depression. Inbreeding increases homozygosity (identical alleles are brought together). Therefore, recessive alleles that were masked by dominant alleles (assuming recessive alleles are less favorable than dominant ones) in parents are expressed. Genetic drift is always present in breeding populations due to the reduction in population size after selection. Inbreeding depression as a result of genetic drift, however, can be modified by increasing effective population sizes or acquiring new genetic materials. Small effective population sizes seem to have been responsible for limited selection responses (Smith, 1983; Tanner and Smith, 1987; Helms et al., 1989; Eyherabide and Hallauer, 1991; Keeratinijacal and Lamkey, 1993; Weyhrich et al., 1998b), especially in early cycles where frequencies of desired alleles were low. However, large improvements with reduced inbreeding depression have been associated with inbred progeny selection in US domestic backgrounds (Oyervides-García and Hallauer, 1986; Tanner and Smith, 1987; Rodriguez and Hallauer, 1988; Carena and Hallauer, 2001), especially when effective population sizes were over 20.

The amount of inevitable inbreeding that occurs in a wholly random mating population is determined by the number $N$ of unrelated individuals in the population, whose gametes unite at random in every generation. In an isolation planting pollinated by wind movement of male gametes, a small amount of self-fertilization can occur, depending on population size. Controlled hand pollinations in a breeding nursery prevent self-fertilization, but inbreeding occurs because of limited population size. Li (1976) showed that the probability that a gamete will unite with a gamete from the same individual is $1/N$ and from a different individual is $(N - 1)/N$, including the possibility of self-fertilization. Because the probability of a gene identical by descent from full-sib matings is $1/4$ (see Chapter 3 ) and the probability of full-sibs mating together is $2/N$, the proportion of individuals identical by descent becomes $1/(2N)$. Hence

$$F = 1/(2N)$$

If $N = 1$ (self-fertilization) then $F = 0.50$ and if $N = 2$ (crosses made between two individuals), $F = 0.25$, which is the coefficient of inbreeding by full-sibbing. In small populations, probabilities of matings among related individuals are high. It has been common in some recurrent selection studies to intermate 10 individuals where the average expected level of inbreeding would be 5% (1/20). Gordillo and Geiger (2008) have shown alternative ways of maintaining larger effective population sizes. However, some studies have shown no advantage to using larger effective population sizes.
sizes (Weyhrich et al., 1998b, Guzman and Lamkey, 2000). In the next cycle or generation we have two proportions that contribute to the level of inbreeding:

(1) Genes at $1/(2N)$ of the loci identical by descent, and
(2) Proportion of homozygous loci because of inbreeding ($F'$) in the previous generation.

The total inbreeding, as shown by Li (1976) is

$$F = 1/(2N) + [1 - 1/(2N)]F'$$

where $1/(2N)$ is the increment for population size and $[1 - 1/(2N)]F'$ is the inbreeding portion already present in the population. This formula is exact and applies to populations whether they are large or small. The number of possible matings among $N$ individuals, therefore, is $N(N - 1)$. This type of mating is common in recurrent selection studies among the selected individuals used to form the next cycle population. The recurrence relation for monoecious individuals that does not include self-fertilization is

$$F = [1/(2N)][1 + 2(N - 1)F' + F'']$$

where $1/(2N)$ is the portion due to population size and $F'$ and $F''$ the increments of inbreeding present in previous cycles. This expression reflects the number of matings made with the restriction preventing self-fertilization. For example, expected total inbreeding in the third cycle from intermating 10 individuals, using the first formula, is 0.1426, whereas expected total inbreeding using the latter formula is 0.1402. The difference in expected total inbreeding reflects the correction for number of matings. Exclusion of self-fertilization shifts the increment of new inbreeding back to the grandparental ($F''$) generation.

If the increment of inbreeding is $\Delta F$, the relation

$$\Delta F = (F' - F)/(1 - F)$$

shows the increase in homozygosity relative to the heterozygosity that is present. The important feature is the heterozygosity present because once the loci are fixed in a closed population they remain fixed unless outcrossing occurs. The ‘new inbreeding’ or ‘increment’ ($\Delta F$) measures the proportionate increase in the rate of inbreeding (Falconer and Mackay, 1996). The rate of inbreeding ($\Delta F$) and population size ($N$), therefore, have the following relation:

$$\Delta F = 1/(2N)$$

A maize population of $N$ diploid monoecious individuals in Hardy–Weinberg equilibrium would produce $2N$ gametes. If we assume the conditions of equal males
and females mating at random, the offspring population of \(N\) individuals will have the expected distribution of \((pA + qa)^{2N}\), which is the familiar binomial expansion. Gene frequency is expected to be the same, but random sampling from a population can cause random deviations from the expected. If we let \(p' - p\) denote the random deviation of gene frequency in the progeny population from the preceding population, the variance of \(p\) becomes

\[
\sigma_p^2 = \frac{[p(1-p)]}{(2N)}
\]

Since \(\Delta F = 1/(2N)\), the increment of the inbreeding coefficient rather than population size can be used to estimate the variance of gene frequency changes:

\[
\sigma_p^2 = \frac{[p(1-p)]}{\Delta F}
\]

Li (1976) used these relations to determine the effective size of a breeding population based on the results of Wright (1931). The rate of decrease in heterozygosis for \(N\) dioecious individuals composed of unequal numbers of males (m) and females (f) was

\[
\frac{1}{(8N_m)} + \frac{1}{(8N_f)}
\]

For equal numbers of males and females, the relation was approximately \(1/(2N)\). If we consider an ideal maize population of \(N\) individuals having equal numbers of males and females mating at random, it was shown that the variance of gene frequency was

\[
[p(1-p)]/2N \text{ and the rate of inbreeding was } 1/(2N)
\]

The effective size of a breeding group of individuals is the actual breeding size compared with the ideal population. Effective size (\(N_e\)) of a population when equal numbers of males and females contribute the same number of gametes to the next generation is

\[
N_e = 2N, \text{ or twice the number of } N \text{ individuals recombined}
\]

With unequal numbers of females and males, the rate of decrease of heterozygosis is

\[
\frac{1}{(8N_m)} + \frac{1}{(8N_f)}, \text{ which becomes } \frac{[(N_m + N_f)]}{(8N_mN_f)} \text{ or } 1/(2N_e)
\]

Rearranging, the effective size of the breeding population is

\[
N_e = \frac{4N_mN_f}{(N_m + N_f)}
\]
Inbreeding which reduces to $2N$ for equal numbers of males and females. Male and female individuals are equally used in most maize populations, but if unequal males and females are used, effective size depends much more on the sex that is fewer in number. For example, for 10 males and 100 females, $N_e = 4(10)(100)/110 = 36.4$. For comparison, for 10 males and 10 females equally mated, $N_e = 4(10)(10)/20 = 20$, which is the same as $N_e = 2N$.

In some recurrent selection programs, progenies that have some level of inbreeding are intermated. The effect of inbreeding is to reduce effective size, as shown by Sprague and Eberhart (1977):

$$N_e = \frac{2N}{1 + F_P}$$

where $F_P$ is the coefficient of inbreeding of the parental plants of the lines being recombined. If the recurrent selection is based on $S_2$ performance itself, one has a choice of recombining remnant $S_1$ or $S_2$ seed. If remnant $S_1$ seed is used, $F_P = 0$ because the gametic array of the $S_1$ lines is equivalent to an $S_0$ plant. Therefore, $N_e$ would be 40 if 20 $S_1$ lines are recombined. If remnant $S_2$ seed is used, then $N_e$ is reduced to 26.7. Because of the reduction in $N_e$ it would be an advantage to use $S_1$ rather than $S_2$ seed for recombination. Another common breeding technique is to recombine elite inbred-lines with above average general combining ability in one or more traits to form synthetic varieties for improvement. For this case, $N_e$ would be equal to the number of lines included. For instance, 16 lines recombined would be the effective size of the population (e.g., the case for BSSS). The intercrossed lines would be random mated but $F$ for $1/(2N)$ is 0.0319 unless some outcrossing, intentional or unintentional, is permitted.

Importance of effective size and associated effects of inbreeding with recurrent selection programs was emphasized by Sprague and Eberhart (1977). It is relatively easy to intercross and random mate sizable populations of maize, but choices have to be made relative to selection intensity and numbers of progenies tested (Robertson, 1960; Baker and Curnow, 1969; Rawlings, 1970). Although the number of individuals selected for recombination can cause some problems in handling, usually the testing phase of the program imposes restraints on the number of individuals recombined. The formulas for inbreeding and effective size are for ideal situations, which rarely occur. Individuals intermated in recurrent selection programs probably do not contribute equally to the next generations, but their contributions are not known unless pedigrees are maintained throughout the selection program. Table 9.2, adapted from Sprague and Eberhart (1977), shows expected levels of inbreeding of different cycles of recurrent selection for different effective size of populations.

It is obvious from Table 9.2 that effective size of population has, in theory, a pronounced effect on expected level of inbreeding and that cumulative effects of inbreeding with cycles of selection are dramatic. If $N = 10$, expected level of inbreeding is 0.22 after five cycles of selection, compared with 0.12 if $N = 20$. As
Table 9.2 Expected levels of inbreeding for the different cycle populations with varying effective size of population

<table>
<thead>
<tr>
<th>Selection cycle</th>
<th>( F = 0 )</th>
<th>( F = 0.5 )</th>
<th>( F = 0 )</th>
<th>( F = 0.5 )</th>
<th>( F = 0 )</th>
<th>( F = 0.5 )</th>
<th>( F = 0 )</th>
<th>( F = 0.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20a</td>
<td>13.3(^a)</td>
<td>30</td>
<td>20</td>
<td>40</td>
<td>26.7</td>
<td>50</td>
<td>33.3</td>
<td>60</td>
</tr>
<tr>
<td>1</td>
<td>0.05</td>
<td>0.08</td>
<td>0.03</td>
<td>0.05</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>0.14</td>
<td>0.06</td>
<td>0.10</td>
<td>0.05</td>
<td>0.07</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.14</td>
<td>0.20</td>
<td>0.09</td>
<td>0.14</td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>0.18</td>
<td>0.25</td>
<td>0.12</td>
<td>0.18</td>
<td>0.10</td>
<td>0.14</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>5</td>
<td>0.22</td>
<td>0.31</td>
<td>0.15</td>
<td>0.22</td>
<td>0.12</td>
<td>0.17</td>
<td>0.10</td>
<td>0.14</td>
</tr>
<tr>
<td>6</td>
<td>0.25</td>
<td>0.35</td>
<td>0.18</td>
<td>0.26</td>
<td>0.14</td>
<td>0.20</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>7</td>
<td>0.29</td>
<td>0.40</td>
<td>0.20</td>
<td>0.29</td>
<td>0.16</td>
<td>0.23</td>
<td>0.13</td>
<td>0.09</td>
</tr>
<tr>
<td>8</td>
<td>0.32</td>
<td>0.44</td>
<td>0.23</td>
<td>0.32</td>
<td>0.18</td>
<td>0.26</td>
<td>0.15</td>
<td>0.21</td>
</tr>
<tr>
<td>9</td>
<td>0.35</td>
<td>0.48</td>
<td>0.26</td>
<td>0.36</td>
<td>0.20</td>
<td>0.28</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>10</td>
<td>0.39</td>
<td>0.52</td>
<td>0.28</td>
<td>0.39</td>
<td>0.22</td>
<td>0.31</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>20</td>
<td>0.62</td>
<td>0.76</td>
<td>0.48</td>
<td>0.62</td>
<td>0.39</td>
<td>0.52</td>
<td>0.33</td>
<td>0.45</td>
</tr>
<tr>
<td>40</td>
<td>0.86</td>
<td>0.94</td>
<td>0.73</td>
<td>0.86</td>
<td>0.63</td>
<td>0.77</td>
<td>0.55</td>
<td>0.69</td>
</tr>
</tbody>
</table>

\(^a\) \( N_e = 2(N/(1 + F)) \), where \( N \) is the number of lines recombined and \( F \) is the inbreeding of the lines recombined; e.g., \( F = 0.5 \) for \( S_2 \) lines.

a consequence, the expected level of inbreeding is 47.8% less for \( N = 20 \) than for \( N = 10 \) after five cycles of selection. After 40 cycles of selection for \( N = 10 \) and \( N = 20 \), expected level of inbreeding is only 26.8% less for \( N = 20 \). Use of \( S_2 \) lines for recombination also shows that expected levels of inbreeding are increased considerably. Use of individuals (or lines) that have some inbreeding reduces effective size of the population. For instance, for \( N = 10 \), \( N_e \) is reduced from 20 to 13.3. Hence expected level of inbreeding in the first cycle is increased by 60%, from 0.05 to 0.08.

Because inbreeding in successive cycles depends on inbreeding in previous cycles, expected level of inbreeding from use of inbred individuals will always be greater than that from use of non-inbred parents. Additional recombination by random mating will not change expected level of fixation as loci that are fixed will remain so unless mutation or outcrossing occurs. To reduce expected level of inbreeding, one alternative would be to use remnant \( S_1 \) seed of the progenitors of the \( S_2 \) progenies that were tested in a recurrent selection program. In all instances, expected level of inbreeding becomes an important factor in long-term selection programs. Inbreeding can be reduced by use of either a larger number of individuals (which either reduces selection intensity or increases number of progenies tested) or individuals that are not inbred for recombination. It is not unreasonable to expect an individual conducting a recurrent selection based on half-sib, full-sib, \( S_1 \), or \( S_2 \) progeny evaluation to complete 10 cycles of selection. If non-inbred individuals are recombined, expected level of inbreeding would be reduced from 0.39 to 0.18 if \( N = 25 \) rather than 10. But again, expected vs. observed level of inbreeding in recurrent selection programs varies according to population undergoing selection, trait under selection, and stage of improvement.
9.5 Estimates of Inbreeding Depression

There are major differences among species for the amount of inbreeding depression that is expressed. This is related with the amount of dominance we encounter in each crop. For instance, in self-pollinated crops inbreeding depression is minimal and homozygous cultivars are used. On the other hand, in some cross-pollinated species inbreeding depression can be severe (e.g., alfalfa) or tolerable where homozygous genotypes can be produced but their performance is reduced (e.g., maize, sunflower). Therefore, inbreeding is not the cause of inbreeding depression.

Different methods have been used to report inbreeding depression. The mean of a non-inbred population is equal to \((p - q)a + 2pqd\) and the mean of a population with an arbitrary level of inbreeding is equal to \((p - q)a + 2pq(1 - F)d\) (see Chapter 2). Therefore, inbreeding depression in absolute values can be defined as \(2pqdF\). Four of the possible methods include the following:

1. \(\bar{X}_0 - \bar{X}_F = 2pqdF\), which is the genetic expectation and expressed in absolute units of the traits measured;
2. \((\bar{X}_0 - \bar{X}_F)/\bar{X}_0 \times 100 = 2pqdF/[(p - q)a + 2pqd)] \times 100\), which is the percentage of inbreeding depression;
3. \((\bar{X}_0 - \bar{X}_F)/F = 2pqd\), which is the genetic expectation per 1% increase in \(F\); and
4. \(b = \) regression of observed means for different levels of inbreeding per 1% increase in homozygosity.

When loci combine additively and with directional dominance, the change in mean with inbreeding should be directly proportional to the coefficient of inbreeding. If epistatic effects are present, there may be a curvilinear change in the population mean with inbreeding. In maize most estimates suggest additive effects primarily determine the means of traits because a linear relation is usually observed between the changes in mean performance and level of inbreeding. The rate of inbreeding depression differs among traits, and the duration of inbreeding depression varies among traits. Grain yield in maize, for example, usually exhibits a decreasing trend from the non-inbred throughout the generations of inbreeding (Fig. 9.1). There is a close relationship between the coefficient of inbreeding and the degree of inbreeding depression and BSSS suffers a reduction of approximately 0.5 q/ha for each 1% homozygosis (see later in the chapter). There is often a linear relationship between percentage of homozygosity and performance of quantitative traits. Plant height and days to flower are examples of traits that tend to stabilize after 3–5 generations of inbreeding.

Although a tremendous number of self-pollinations have been made in maize, estimates of the inbreeding depression for different traits are surprisingly few. From the earliest studies in maize (Shamel, 1905; East, 1908; Shull, 1908) effects of inbreeding were obvious. With increasing homozygosity

(1) Vigor and productiveness were reduced and traits became fixed.
(2) Differences among lines increased whereas variability within lines decreased.
Effects of inbreeding were interpreted on the basis of Mendelian genetics (Shull, 1908) because of fixation of alleles with increased homozygosity. Means and frequency distributions (Jones, 1918; Shull, 1952) usually were given, but rates of inbreeding depression were not given. In most instances too few lines were included and the observations were made in different years for different generations. Estimates would have been biased by environmental effects in the different years and the errors of estimation large. For the limited number of lines included, detailed data were collected and observations recorded for each line in each generation. Information was then interpreted correctly relative to the expected level of homozygosity by self-fertilization on the basis of Mendelian segregation of genetic factors.

The classic study of the effects of self-fertilization in maize was reported by Jones (1918, 1939). The study originally started with 12 self-fertilizations in the open-pollinated variety Chester’s Leaming in 1904 and the first detailed report was given by East and Hayes (1912). Jones (1918) reported on the first 11 generations of self-fertilization for four inbred-lines (two of which were sub-strains of the original cross) for yield and plant height. Measurements were made in each year of self-fertilization. A gradual reduction in yield and plant height with self-fertilization occurred, but environmental effects for estimating rate of inbreeding depression were obvious. For example, yield of line 1–6–1–3 was 49.2 q/ha in generation 8 vs. 17.3 and 15.9 q/ha in generations 5 and 9, respectively. Self-fertilization for three lines was continued for 30 generations and the results were given by Jones (1939). Because of the disturbing environmental effects in each generation, he grouped the measurements for each five-generation segment to average the seasonal fluctuations (Table 9.3). A decreasing trend for yield occurred throughout the 30 generations of self-fertilization, but ear height showed no reduction after five generations. Linear regressions for the decrease in means with increased homozygosity were similar for each line for plant height and yield. On the average, plant height was reduced
Table 9.3  Effects of 30 generations of self-fertilization on plant height and grain yield for three lines of maize (adapted from Jones, 1939)

<table>
<thead>
<tr>
<th>Generations selfed</th>
<th>Inbreds-lines (1–6, 1–7, 1–9)</th>
<th>Average % Change</th>
<th>Inbreds-lines (1–6, 1–7, 1–9)</th>
<th>Average % Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>297.2 297.2 297.2 297.2</td>
<td>—</td>
<td>50.6 50.6 50.6 50.6</td>
<td>—</td>
</tr>
<tr>
<td>1–5</td>
<td>221.0 205.7 195.6 207.5</td>
<td>30</td>
<td>40.0 31.9 25.6 32.5</td>
<td>36</td>
</tr>
<tr>
<td>6–10</td>
<td>246.4 213.4 208.3 222.8</td>
<td>25</td>
<td>28.1 22.5 21.2 23.9</td>
<td>53</td>
</tr>
<tr>
<td>11–15</td>
<td>246.4 213.4 210.8 223.5</td>
<td>25</td>
<td>23.8 21.2 16.2 20.4</td>
<td>60</td>
</tr>
<tr>
<td>16–20</td>
<td>223.5 215.9 190.5 210.0</td>
<td>29</td>
<td>13.8 15.0 8.8 12.5</td>
<td>75</td>
</tr>
<tr>
<td>21–25</td>
<td>205.7 190.5 180.3 192.3</td>
<td>35</td>
<td>12.5 13.1 8.1 11.2</td>
<td>78</td>
</tr>
<tr>
<td>26–30</td>
<td>233.5 203.2 195.6 210.8</td>
<td>29</td>
<td>15.0 11.2 5.6 10.6</td>
<td>79</td>
</tr>
<tr>
<td>h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>−8.7  −11.1  −12.6  −10.8</td>
<td>—</td>
<td>−6.3  −5.8  −6.5  −6.2</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plant height (cm)</th>
<th>Yield (q/ha)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Generations selfed</th>
<th>Inbreds-lines (1–6, 1–7, and 1–9)</th>
<th>Average % Change</th>
<th>Inbreds-lines (1–6, 1–7, and 1–9)</th>
<th>Average % Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>297.2 297.2 297.2 297.2</td>
<td>—</td>
<td>50.6 50.6 50.6 50.6</td>
<td>—</td>
</tr>
<tr>
<td>1–5</td>
<td>221.0 205.7 195.6 207.5</td>
<td>30</td>
<td>40.0 31.9 25.6 32.5</td>
<td>36</td>
</tr>
<tr>
<td>6–10</td>
<td>246.4 213.4 208.3 222.8</td>
<td>25</td>
<td>28.1 22.5 21.2 23.9</td>
<td>53</td>
</tr>
<tr>
<td>11–15</td>
<td>246.4 213.4 210.8 223.5</td>
<td>25</td>
<td>23.8 21.2 16.2 20.4</td>
<td>60</td>
</tr>
<tr>
<td>16–20</td>
<td>223.5 215.9 190.5 210.0</td>
<td>29</td>
<td>13.8 15.0 8.8 12.5</td>
<td>75</td>
</tr>
<tr>
<td>21–25</td>
<td>205.7 190.5 180.3 192.3</td>
<td>35</td>
<td>12.5 13.1 8.1 11.2</td>
<td>78</td>
</tr>
<tr>
<td>26–30</td>
<td>233.5 203.2 195.6 210.8</td>
<td>29</td>
<td>15.0 11.2 5.6 10.6</td>
<td>79</td>
</tr>
<tr>
<td>h&lt;sup&gt;a&lt;/sup&gt;</td>
<td>−8.7  −11.1  −12.6  −10.8</td>
<td>—</td>
<td>−6.3  −5.8  −6.5  −6.2</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>a</sup> Linear regression of trait on seven groups of self-fertilization

29% and yield reduced 79% after 30 generations of self-fertilization. Effects of self-fertilization for the two traits were quite different. After five generations plant height was reduced 30% with little change thereafter. On the other hand, yield continued to decrease with self-fertilization and was reduced 75% by generation 20. Jones (1939) discusses possible reasons for the differences in inbreeding depression for the two traits. Because the level of heterozygosity is reduced 50% on the average with each self-fertilization, the more rapid attainment of stability of means for plant height may be due to less complexity in its inheritance than for yield. If fewer genetic factors influence the expression for plant height, homozygosity would be approached more rapidly than for a more complex trait, particularly in grain yield. Grain yield is the ultimate expression of the productivity of a genotype that includes expression of physiological mechanisms throughout the growing season. Average estimates of heritability (e.g., individual plant or a progeny basis, given in Chapter 5) show that the estimate of heritability for plant height is about three times greater than for yield.

Jones (1939) produced and tested crosses of the sub-strains of the 1–6 and 1–7 lines (Table 9.3) and found on evidence that the F<sub>1</sub> significantly differed from either parent.

An illustration of the approach to homozygosity is given in Fig. 9.2 for different numbers of genetic factors. Loss of heterozygosity, on the average, is less as the number of genetic factors increases.

Expected relative heterozygosity is given for the same locus in all individuals in the population, or for all loci within the same individual. As the number of loci affecting a trait is increased, loss of heterozygosity for the population is not as rapid as for say one factor, where the heterozygosity is reduced 50% after one generation of self-fertilization. Accurate numbers of genetic factors affecting most traits considered quantitative in inheritance are still unknown (e.g., polygenes, QTL), but Jones’s
Fig. 9.2 Percentage of heterozygous individuals in the population for each generation of self-fertilization when 1, 5, 10, and 15 allelic pairs are considered.

Data certainly suggest fewer factors affecting plant height than grain yield. From Fig. 9.2 and Allard (1960) the number of factors affecting yield must be greater than 15 because the approach to homozygosity is slower than shown for 15 factors.

Estimates of inbreeding depression in maize studies have shown the relative rates for different plant and ear traits. In comparison to the study reported by Jones (1939) and others, these studies included a greater number of progenies and progeny sets compared in the same environments by use of experimental designs. Remnant seeds of different generations of inbreeding and the original population were compared in the same experiments to obtain estimates of environmental effects and experimental error. Estimates of rates of inbreeding depression were calculated by regression of means for different generations of inbreeding on expected level of homozygosity. Estimates of inbreeding depression for several traits of maize with at least eight estimates are given in Table 9.4 for 11 populations, but five of these are for advanced cycles of six populations.

The precision of average estimates for the different traits is not the same because different numbers of estimates were available, ranging from eight for ear number to 19 for yield. Estimates of inbreeding depression for yield are given on a per area basis (q/ha) and per plant basis (g/plant). For yield, we can expect 0.5 q/ha (1.2 g/plant) decrease for each 1% increase in homozygosity. On a per plant basis, estimates of inbreeding depression for yield were similar for all populations, ranging from −0.82 (Synthetic O.P. by full-sibbing) to −1.44 (BSSSCO) g/plant. Although
Table 9.4  Inbreeding depression showing the change in phenotypic means per 1% increase of the coefficient of inbreeding for different maize populations

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Yield (g/ha)</th>
<th>Height (cm)</th>
<th>Kernel weight (grams)</th>
<th>Ear number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Plant</td>
<td>Ear</td>
<td>Days to flower</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sing et al. (1967)</td>
<td>Indian Chief</td>
<td>−0.2018</td>
<td>−0.2400</td>
<td>−0.0225</td>
<td>−0.0004</td>
</tr>
<tr>
<td></td>
<td>Jarvis</td>
<td>−0.2360</td>
<td>−0.3300</td>
<td>+0.0410</td>
<td>−0.0004</td>
</tr>
<tr>
<td>Genter (1971)</td>
<td>BSSSC0</td>
<td>−0.7820</td>
<td>−0.8438</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BSSS(HT)C7</td>
<td>−0.5560</td>
<td>−0.2400</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VaCBS(C0)</td>
<td>−0.7760</td>
<td>−0.2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VaCBS(S)C4</td>
<td>−0.6510</td>
<td>−0.1800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harris et al. (1972)</td>
<td>NHG</td>
<td>−0.7210</td>
<td>−0.2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NHG(M)C9</td>
<td>−0.7172</td>
<td>−0.1800</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NHG(I)C9</td>
<td>−0.6232</td>
<td>−0.1600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hallauer and Sears (1973)</td>
<td>BSSSC0</td>
<td>−0.4490</td>
<td>−0.2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cornelius and Dudley (1974)</td>
<td>Synthetic O.P.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sibbing</td>
<td>−0.3855</td>
<td>−0.5805</td>
<td>−0.0427</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>−0.3797</td>
<td>−0.5760</td>
<td>+0.0441</td>
<td></td>
</tr>
<tr>
<td>Good and Hallauer (1977)</td>
<td>BSSSC0</td>
<td>−0.3976</td>
<td>−0.3955</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sibbing</td>
<td>−0.4628</td>
<td>−0.5760</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>−0.4511</td>
<td>−0.5760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice and Dudley (1974)</td>
<td>Autotetraploids</td>
<td>−1.8426</td>
<td>−1.8426</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sibbing</td>
<td>−1.5360</td>
<td>−1.5360</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full-sibbing</td>
<td>−1.5360</td>
<td>−1.5360</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>−0.5104</td>
<td>−0.4398</td>
<td>+0.0499</td>
<td>−0.1163</td>
</tr>
<tr>
<td>Levings et al. (1967)</td>
<td>Autotetraploids</td>
<td>−1.3210</td>
<td>−1.3210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice and Dudley (1974)</td>
<td>Sibbing</td>
<td>−1.3210</td>
<td>−1.3210</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full-sibbing</td>
<td>−1.3210</td>
<td>−1.3210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
estimates of inbreeding depression for yield are similar, they were obtained in different populations evaluated across different environments with theoretical levels of homozygosity attained by different systems of inbreeding. Sing et al. (1967) obtained their levels of inbreeding from the pedigrees of double–double crosses. Genter (1971) and Harris et al. (1972) obtained their estimates from S1 lines. Hallauer and Sears (1973), Cornelius and Dudley (1974), and Good and Hallauer (1977) obtained their estimates by comparing different generations of inbreeding attained by either self-fertilization or sibbing. Therefore, in spite of differences in populations, environments, and procedures for attaining expected levels of homozygosity, estimates of the linear decrease in yield per plant for each 1% increase in inbreeding were nearly the same. Differences in yield loss with inbreeding on a per area basis are greater among the studies than on a per plant basis because different plant stand densities were used in the different studies. If linear rates of decrease per plant were adjusted to a constant stand, the rates of decrease in yield with increasing homozygosity would be more similar. Response in yield with inbreeding corroborates observations reported from earliest studies in maize in which yield or productivity decreased in all instances.

Estimates of inbreeding depression were negative for all other traits except days to flower. Six estimates for barrenness percentage were also positive (Good and Hallauer, 1977). These expressions of inbreeding also agree with previous reported effects for reduced vigor and plant size, later flowering, and increased barrenness. Days to flower, for example, had positive estimates in all instances, ranging from +0.02 to +0.09 for an average of +0.05. Estimates obtained by selfing or sibbing also were similar. Inbreeding is expected to reduce yield and components of yield, reduce plant size, and increase time to flowering and incidence of barrenness confirming previous observations.

Expected changes from inbreeding of the traits listed in Table 9.4 are shown in Table 9.5 for five levels of homozygosity. Average estimates for each 1% increase in level of homozygosity (Table 9.4) were used to calculate the other four levels of homozygosity. The calculated means in Table 9.5 assume that linear regression is a valid mathematical description of the average performance of a population for different levels of inbreeding. Assuming a linear relation, we can expect on the average that grain yield is reduced 115.2 g/plant from the non-inbred generation to 100% homozygosity, or on a per unit basis we can expect a 51 q/ha decrease in yield. The population of inbred-lines on the average is expected to be 44 cm shorter, have 31 cm lower ear placement, and be about 5 days later in maturity.

Percentage of decrease by self-fertilization for an unselected set of inbred-lines developed from Iowa Stiff Stalk Synthetic was reported by Good and Hallauer (1977). The relative decreases from the original non-inbred population to the S8 generation (eight generations of self-fertilization) are shown in Table 9.6. The relative decrease from the S0 to the S8 generation was greatest for yield, which was reduced 68%. There was a steady decline in yield throughout the range of inbreeding, which agrees with Jones’s (1939) data. Plant height, the other trait studied by Jones, was reduced 22% by the S3 generation with only 2% additional reduction for the five additional generations of self-fertilization, which also agrees with Jones. Ear height
Table 9.5  Inbreeding depression expected from the estimates obtained from maize populations (Table 9.4) for five levels of homozygosity

<table>
<thead>
<tr>
<th>Trait</th>
<th>Level of homozygosity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Yield (q/ha)</td>
<td>−0.5104</td>
</tr>
<tr>
<td>Yield (g/plant)</td>
<td>−1.1517</td>
</tr>
<tr>
<td>Plant height (cm)</td>
<td>−0.4398</td>
</tr>
<tr>
<td>Ear height (cm)</td>
<td>−0.3124</td>
</tr>
<tr>
<td>Days to flower (no.)</td>
<td>+0.0499</td>
</tr>
<tr>
<td>Ear number</td>
<td>−0.0004</td>
</tr>
<tr>
<td>Grain moisture (%)</td>
<td>−0.0046&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Shelling (%)</td>
<td>−0.0720&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ear diameter, (mm)</td>
<td>−0.0926&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Earl length (cm)</td>
<td>−0.3518&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Kernel-row (no.)</td>
<td>−0.0180&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Kernel depth (mm)</td>
<td>−0.0642&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Kernel weight (g)</td>
<td>−0.1163</td>
</tr>
<tr>
<td>Oil (%)</td>
<td>−0.0034&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Cob diameter (mm)</td>
<td>−0.0355&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Leaf width (mm)</td>
<td>−0.1340&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Avg. barrenness</td>
<td>+0.0009&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Stand (no.)</td>
<td>−0.0792&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Average based on less than eight estimates, more estimates are encouraged to validate expected inbreeding depression values
Table 9.6  Observed means by selfing in Iowa Stiff Stalk Synthetic (BSSS) and percentage changes in traits relative to original non-inbred population

<table>
<thead>
<tr>
<th>Generation</th>
<th>Yield, g</th>
<th>Plant height, cm</th>
<th>Ear height, cm</th>
<th>Days to flower</th>
<th>Ear number</th>
<th>Ear diameter, mm</th>
<th>Ear length, mm</th>
<th>Kernel depth, mm</th>
<th>Kernel weight, g</th>
<th>Cob diameter, mm</th>
<th>Barrenness</th>
<th>Stand</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC&lt;a&gt;</td>
<td>172.3</td>
<td>102</td>
<td>192.1</td>
<td>99</td>
<td>95.7</td>
<td>97.26.2</td>
<td>98</td>
<td>0.12</td>
<td>120</td>
<td>48</td>
<td>100</td>
<td>185</td>
</tr>
<tr>
<td>DC&lt;a&gt;</td>
<td>165.5</td>
<td>98</td>
<td>195.4</td>
<td>101</td>
<td>97.9</td>
<td>99.26.9</td>
<td>101</td>
<td>0.08</td>
<td>80</td>
<td>48</td>
<td>100</td>
<td>184</td>
</tr>
<tr>
<td>BSSS S₀</td>
<td>168.6</td>
<td>100</td>
<td>194.1</td>
<td>100</td>
<td>98.9</td>
<td>100.26.7</td>
<td>100</td>
<td>0.10</td>
<td>100</td>
<td>48</td>
<td>100</td>
<td>180</td>
</tr>
<tr>
<td>S₁</td>
<td>112.1</td>
<td>66</td>
<td>175.5</td>
<td>90</td>
<td>84.1</td>
<td>85.30.4</td>
<td>114</td>
<td>0.12</td>
<td>120</td>
<td>45</td>
<td>94</td>
<td>166</td>
</tr>
<tr>
<td>S₂</td>
<td>87.0</td>
<td>52</td>
<td>162.0</td>
<td>83</td>
<td>73.9</td>
<td>75.30.3</td>
<td>113</td>
<td>0.13</td>
<td>130</td>
<td>43</td>
<td>90</td>
<td>155</td>
</tr>
<tr>
<td>S₃</td>
<td>75.1</td>
<td>44</td>
<td>152.4</td>
<td>78</td>
<td>69.6</td>
<td>70.31.1</td>
<td>116</td>
<td>0.13</td>
<td>130</td>
<td>42</td>
<td>88</td>
<td>150</td>
</tr>
<tr>
<td>S₄</td>
<td>64.9</td>
<td>38</td>
<td>150.5</td>
<td>78</td>
<td>67.5</td>
<td>68.32.1</td>
<td>120</td>
<td>0.13</td>
<td>130</td>
<td>41</td>
<td>85</td>
<td>143</td>
</tr>
<tr>
<td>S₅</td>
<td>61.2</td>
<td>36</td>
<td>148.1</td>
<td>76</td>
<td>65.7</td>
<td>66.32.3</td>
<td>121</td>
<td>0.12</td>
<td>120</td>
<td>41</td>
<td>85</td>
<td>143</td>
</tr>
<tr>
<td>S₆</td>
<td>54.4</td>
<td>32</td>
<td>147.9</td>
<td>76</td>
<td>64.3</td>
<td>65.33.5</td>
<td>125</td>
<td>0.13</td>
<td>130</td>
<td>40</td>
<td>83</td>
<td>140</td>
</tr>
<tr>
<td>(SE) X</td>
<td>2.8</td>
<td>2.0</td>
<td>1.6</td>
<td>0.6</td>
<td>0.01</td>
<td>0.2</td>
<td>1.4</td>
<td>0.2</td>
<td>0.8</td>
<td>0.2</td>
<td>0.01</td>
<td>0.8</td>
</tr>
</tbody>
</table>

<br/><br/><sup>a</sup>Bulk of single crosses (SC) and double crosses (DC) among S₇ lines included as checks
had a greater relative reduction with inbreeding, with a decreasing trend in all generations. The change was only 5% after the S₃ generation vs. 30% for the first three generations of self-fertilization. Days to flower increased 25% from S₀ to S₈ and increased throughout the range of self-fertilization, but the greatest one-generation increase occurred from S₀ to S₁. Of yield components, kernel depth exhibited the greatest decrease with inbreeding (37%) and occurred throughout the generations of self-fertilization. Although none of the components of yield was affected by inbreeding to the same extent as yield itself, the cumulative effects of inbreeding for all traits affected yield.

Good and Hallauer (1977) conducted a comprehensive study estimating and comparing rates of inbreeding depression for unselected lines developed by three methods of inbreeding: self-fertilization, full-sibbing, and full-sibbing followed by self-fertilization. Each of the three series of lines was isolated from Iowa Stiff Stalk Synthetic with no intentional selection during the generations of inbreeding. Originally 250 S₀ plant were self-fertilized and 243 pairs of S₀ plants were sib-mated. Two separate samples of S₀ plants were used to develop the self and sib series of lines. The sib series was sib-mated for five generations at which time two progeny rows were grown. Sib-mating was continued in one row and self-fertilization was initiated in the second row. Bulk entries of each of the three series of inbred-lines for the generations of inbreeding were evaluated in five-replicate experiments at nine Iowa environments. Rates of inbreeding depression were determined for each series of lines (Good and Hallauer, 1977) and are included in Table 9.4. Although some statistically significant differences occurred among the linear regression coefficients, the differences were small (Table 9.4). Linear regressions among the three series of lines were significantly different for plant height, cob diameter, yield, 300-kernel weight, stand, and barrenness but, for example, linear regression estimates for yield (−1.15, −1.12, and −1.13 for self-fertilization, full-sibbing, and full-sibbing, and self-fertilization, respectively) did not have much practical significance. Relative rates of inbreeding depression (Fig. 9.3) emphasize the similarity of the decrease in yield by the three methods of inbreeding.

There is little indication that inbreeding by full-sibbing rather than selfing will result in more vigorous lines, although the rate of inbreeding depression for yield is slightly less for the two instances that self-fertilization and full-sibbing were directly compared (Cornelius and Dudley, 1974; Good and Hallauer, 1977, Table 9.4). Good and Hallauer (1977) found that the linear model accounted for more than 99% of the variation for yield.

Good and Hallauer (1977) also compared the means of the traits for the three methods of inbreeding at comparable levels of homozygosity. The theoretical levels of homozygosity were not exactly the same, but they were similar enough for comparative purposes (Table 9.7). In very few instances were the means at similar levels of homozygosity significantly different. Probably the most important difference was for yield at the greatest level of homozygosity. Inbred-lines developed by full-sibbing for 10 generations and selfing for three generations had an average yield that was significantly greater than the lines developed by self-fertilization (25.7 vs.
Fig. 9.3 Inbreeding depression of yield in Iowa Stiff Stalk Synthetic by selfing (S), full-sibbing (FS), and full-sibbing followed by three generations of selfing (FSS).

22.0 q/ha). It is doubtful that the yield advantage is great enough to warrant the five additional generations of inbreeding required to attain the same level of homozygosity. The data in Table 9.7 substantiate the theoretical levels of homozygosity for the three methods of inbreeding. It seems that for unselected lines the effects of homozygosity are the same whatever method of inbreeding is used to attain it. Empirical data agree with theoretical expectations for the different methods. Except for the belief in the effectiveness of visual selection for milder forms of inbreeding, it does not seem that other forms can be recommended. Effects of inbreeding are the same whatever method is used to attain a given level. For instance, 50% homozygosity can be attained by one generation of self-fertilization or three generations of full-sibbing, but in only one instance (ear diameter) were the generation means significantly different. At expected homozygosity levels of 99%, the generations for the three methods of inbreeding significantly differed in five instances. In all except one (barrenness), the means of the lines developed by 10 generations of sib-mating and three generations of self-fertilization were greater than those obtained by eight generations of self-fertilization. Although significant, the differences for the five traits were relatively small in all instances.
Table 9.7  Comparisons between generation means with comparable levels of expected homozygosity for lines developed by three methods of inbreeding

<table>
<thead>
<tr>
<th>Generation†</th>
<th>Expected homozygosity (%)</th>
<th>Yield (q/ha)</th>
<th>(g/plant)</th>
<th>Plant height (cm)</th>
<th>Ear height (cm)</th>
<th>Days to flower§</th>
<th>Ear no.</th>
<th>Ear diam. (mm)</th>
<th>Ear len. (mm)</th>
<th>Kernel depth (mm)</th>
<th>Kernel weight (g)</th>
<th>Cob diameter (mm)</th>
<th>Average barrenness (%)</th>
<th>Stand no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S₀</td>
<td>0.0</td>
<td>68.0</td>
<td>168.6</td>
<td>194.1</td>
<td>98.9</td>
<td>26.7</td>
<td>1.10</td>
<td>48</td>
<td>180</td>
<td>19</td>
<td>77.3</td>
<td>30</td>
<td>0.04</td>
<td>44.3</td>
</tr>
<tr>
<td>S₁</td>
<td>50.0</td>
<td>45.2</td>
<td>112.1</td>
<td>175.5</td>
<td>84.1</td>
<td>30.4</td>
<td>1.12</td>
<td>45ₐ</td>
<td>166</td>
<td>16</td>
<td>71.3</td>
<td>29</td>
<td>0.04</td>
<td>39.1</td>
</tr>
<tr>
<td>FS₃</td>
<td>50.0</td>
<td>45.9</td>
<td>113.5</td>
<td>170.2</td>
<td>81.4</td>
<td>30.6</td>
<td>1.08</td>
<td>46ₐ</td>
<td>164</td>
<td>16</td>
<td>72.7</td>
<td>29</td>
<td>0.04</td>
<td>38.2</td>
</tr>
<tr>
<td>FS₆</td>
<td>73.4</td>
<td>37.7</td>
<td>93.5</td>
<td>159.3</td>
<td>73.5</td>
<td>29.6</td>
<td>1.13</td>
<td>44ₐ</td>
<td>157</td>
<td>15</td>
<td>70.9</td>
<td>29</td>
<td>0.06</td>
<td>38.7</td>
</tr>
<tr>
<td>S₂</td>
<td>75.0</td>
<td>35.1</td>
<td>87.0</td>
<td>162.0</td>
<td>73.9</td>
<td>30.3</td>
<td>1.13</td>
<td>43ₐ</td>
<td>155</td>
<td>15</td>
<td>70.3</td>
<td>28</td>
<td>0.07</td>
<td>37.7</td>
</tr>
<tr>
<td>FS₈</td>
<td>82.6</td>
<td>33.8</td>
<td>83.8</td>
<td>153.6</td>
<td>69.3</td>
<td>30.6</td>
<td>1.12</td>
<td>43</td>
<td>152</td>
<td>14</td>
<td>69.8</td>
<td>28</td>
<td>0.08</td>
<td>35.8</td>
</tr>
<tr>
<td>FS₅–S₁</td>
<td>83.6</td>
<td>31.9</td>
<td>79.0</td>
<td>152.8</td>
<td>69.4</td>
<td>30.5</td>
<td>1.12</td>
<td>43</td>
<td>150</td>
<td>15</td>
<td>68.4</td>
<td>28</td>
<td>0.08</td>
<td>38.3</td>
</tr>
<tr>
<td>S₃</td>
<td>87.5</td>
<td>30.3</td>
<td>75.1</td>
<td>152.4</td>
<td>69.6</td>
<td>31.0</td>
<td>1.13</td>
<td>42</td>
<td>150</td>
<td>14</td>
<td>69.2</td>
<td>28</td>
<td>0.10</td>
<td>35.3</td>
</tr>
<tr>
<td>FS₉</td>
<td>85.9</td>
<td>31.2</td>
<td>77.4</td>
<td>150.2</td>
<td>66.8</td>
<td>30.1</td>
<td>1.13</td>
<td>42</td>
<td>151</td>
<td>14</td>
<td>70.0</td>
<td>28</td>
<td>0.02</td>
<td>36.9</td>
</tr>
<tr>
<td>FS₁₀</td>
<td>88.6</td>
<td>26.7</td>
<td>66.1</td>
<td>148.4</td>
<td>67.6</td>
<td>30.7</td>
<td>1.13</td>
<td>41</td>
<td>145</td>
<td>13</td>
<td>69.8</td>
<td>28</td>
<td>0.13</td>
<td>35.3</td>
</tr>
<tr>
<td>S₄</td>
<td>93.8</td>
<td>26.2</td>
<td>64.9</td>
<td>150.5</td>
<td>67.5</td>
<td>32.1</td>
<td>1.13</td>
<td>41</td>
<td>143</td>
<td>13</td>
<td>69.2</td>
<td>28</td>
<td>0.13</td>
<td>33.0</td>
</tr>
<tr>
<td>FS₅–S₂</td>
<td>91.8</td>
<td>27.4</td>
<td>67.9</td>
<td>147.9</td>
<td>67.1</td>
<td>31.7</td>
<td>1.12</td>
<td>41</td>
<td>146</td>
<td>13</td>
<td>67.5</td>
<td>28</td>
<td>0.08</td>
<td>36.6</td>
</tr>
<tr>
<td>S₆</td>
<td>98.4</td>
<td>24.7</td>
<td>61.2</td>
<td>148.1</td>
<td>65.7</td>
<td>32.3</td>
<td>1.12</td>
<td>41ₐ</td>
<td>143</td>
<td>13</td>
<td>67.5</td>
<td>28</td>
<td>0.12</td>
<td>33.9</td>
</tr>
<tr>
<td>FS₅–S₄</td>
<td>98.0</td>
<td>26.1</td>
<td>64.6</td>
<td>144.2</td>
<td>65.7</td>
<td>31.1</td>
<td>1.12</td>
<td>41ₐ</td>
<td>144</td>
<td>13</td>
<td>68.3</td>
<td>28</td>
<td>0.12</td>
<td>34.6</td>
</tr>
<tr>
<td>FS₁₀–S₂</td>
<td>97.2</td>
<td>24.9</td>
<td>61.7</td>
<td>146.1</td>
<td>63.6</td>
<td>31.5</td>
<td>1.15</td>
<td>40ₐ</td>
<td>141</td>
<td>13</td>
<td>69.4</td>
<td>28</td>
<td>0.12</td>
<td>33.3</td>
</tr>
<tr>
<td>S₈</td>
<td>99.6</td>
<td>22.0</td>
<td>54.4</td>
<td>147.9</td>
<td>64.3</td>
<td>33.5</td>
<td>1.13</td>
<td>40</td>
<td>140</td>
<td>12</td>
<td>65.9</td>
<td>28</td>
<td>0.14</td>
<td>35.0</td>
</tr>
<tr>
<td>FS₁–S₅</td>
<td>99.5</td>
<td>22.8</td>
<td>56.2</td>
<td>142.6</td>
<td>65.3</td>
<td>32.2</td>
<td>1.12</td>
<td>40</td>
<td>138</td>
<td>12</td>
<td>67.5</td>
<td>27</td>
<td>0.09</td>
<td>33.8</td>
</tr>
<tr>
<td>FS₁₀–S₃</td>
<td>98.6</td>
<td>25.7</td>
<td>63.7</td>
<td>144.8</td>
<td>67.0</td>
<td>31.6</td>
<td>1.13</td>
<td>40</td>
<td>142</td>
<td>13</td>
<td>69.6</td>
<td>28</td>
<td>0.12</td>
<td>33.0</td>
</tr>
</tbody>
</table>

†Sᵢ, FSᵢ, and FSᵢ – Sᵢ indicate number of generations of self-fertilization, full-sibling, and full-sibling followed by self-fertilization, respectively

‡Generation means with different letters are significantly (\(P \leq 0.05\)) different. All comparisons without superscripts are not significantly different

§Days after July 1
Predicted linear and quadratic models as well as observed means for yield are shown in Fig. 9.4. There is a slight quadratic trend but the inbreeding depression for yield is essentially linear with increased homozygosity.

The inbreeding study conducted by Good and Hallauer (1977) included 22 levels of homozygosity for the three series of lines developed by self-fertilization, full-sibbing, and combinations of self-fertilization and full-sibbing. Additionally, a level of homozygosity for 12.5% was established from half-sibs. Linear and quadratic regression models were fitted for the means of each trait for the 22 levels of homozygosity. Linear regression coefficients for the linear model were significantly different from zero for all traits. They were positive for days to flower, ears per plant, and barrenness while they were negative for all other traits. Fitting the quadratic model resulted in non-significant linear regressions for ear diameter, ears per plant, and barrenness, and in non-significant quadratic regressions for days to flower, plant and ear heights, cob diameter, stand, and ears per plant. Although quadratic regressions were significant for 6 of the 12 traits, the amount of total variation among the levels of homozygosity accounted for by quadratic regression was small. Linear prediction for yield accounted for 99.1% of total variation.
Wright (1922b) determined for independent loci that the mean performance of a trait is proportional to the decrease in heterozygosis (or increase in homozygosis), regardless of the number of alleles or level of dominance at each locus. A linear decrease would be expected with a decrease in heterozygosis. A curvilinear response is interpreted as evidence of inter-allelic interactions or epistasis. All the inbreeding studies relating mean performance to level of heterozygosity seemed to be adequately described by a genetic model that includes additivity of unlinked locus effects. Linear regression coefficients generally accounted for most of the variation among the generations of inbreeding. In most instances the proportion of the total sums of squares among generations explained by linear regression or percentage of homozygosity $F$ exceeded 90%, usually it was greater than 98–99%. Dominance (intra-locus interactions) is necessary to exhibit inbreeding depression and linear regression measures the net dominance deviation. The absence of any appreciable deviation from regression does not mean that inter-locus interactions (or epistasis) do not exist; it could reflect the net inter-locus interactions caused by cancellation effects. Because there are generally no significant deviations from the linear model, the effects of epistasis on inbreeding depression do not seem to be important. The linear regression model seems to be appropriate for describing average performance of different generations of inbreeding regardless of the system of inbreeding used in maize.

Similar to heterosis, the effects of inbreeding depression (ID) are influenced by allele frequencies, directional dominance, and number of loci (Fig. 9.2). The expression presented by Falconer and Mackay (1996) 

$$\text{ID} = \sum 2pqdF$$

illustrates how estimates of ID can vary among the studies that include different sets of genetic materials. If one assumes two alleles per locus the rate of ID will be at a maximum when allele frequencies are 0.5 and decreases as the allele frequencies approach either 0 or 1. Rodriguez and Hallauer (1988) and Lamkey and Smith (1987) show how estimates of ID can vary among populations and how estimates of ID can change with different methods of selection (Table 9.8). 

In Table 9.8 inter-population selection included either half-sib or full-sib reciprocal recurrent selection and recurrent selection based on testcrosses where direct response to selection is measured in the population crosses. Intra-population recurrent selection included either half-sib or S1–S2 selection, and direct response to selection was measured for the populations themselves. The changes in estimates of ID suggest that either the frequencies of favorable alleles increased with selection or the number of segregating loci had increased with selection, or may be both had occurred. For the Iowa Stiff Stalk Synthetic (BSSS) population, the estimate of ID for the original population was 37.4% because grain yield was reduced 37.4% after one generation of self-pollination. After 10 cycles of reciprocal recurrent selection, the estimate of ID for BSSS(R)C10 was 24.7% or 12.7% less than for BSSSC0. Twelve cycles of selection (eight cycles of half-sib and four cycles of $S_1$ recurrent selection) have been completed in BS13(S)C4; the estimate of ID was 12.7 or 24.7%
### Table 9.8
Estimates of inbreeding depression (ID) for grain yield (%) on maize populations (C0) and after recurrent selection (Ci), primarily for grain yield

<table>
<thead>
<tr>
<th>Population</th>
<th>$S_1/C0$</th>
<th>$S_1/Ci$</th>
<th>Population</th>
<th>$S_1/C0$</th>
<th>$S_1/Ci$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSSC0</td>
<td>37.4 (44.4)$^b$</td>
<td>—</td>
<td>BSSSC0</td>
<td>54.9</td>
<td>—</td>
</tr>
<tr>
<td>BSS(R)C10</td>
<td>15.0 (26.4)</td>
<td>24.7</td>
<td>BSS(H)C8</td>
<td>11.8</td>
<td>30.4</td>
</tr>
<tr>
<td>BS13(S)C4</td>
<td>−14.4 (17.7)</td>
<td>12.7</td>
<td>BSK(S)C8</td>
<td>3.0</td>
<td>26.4</td>
</tr>
<tr>
<td>BSCB1C0</td>
<td>38.3</td>
<td>—</td>
<td>BS12C0</td>
<td>55.9</td>
<td>—</td>
</tr>
<tr>
<td>BSCB1(R)C10</td>
<td>47.7</td>
<td>37.5</td>
<td>BS12(H)C7</td>
<td>−34.7</td>
<td>29.8</td>
</tr>
<tr>
<td>BS10C0</td>
<td>32.1</td>
<td>—</td>
<td>BS2C0</td>
<td>50.3</td>
<td>—</td>
</tr>
<tr>
<td>BS10(FR)C7</td>
<td>27.2</td>
<td>36.1</td>
<td>BS2(S)C4</td>
<td>12.7</td>
<td>29.7</td>
</tr>
<tr>
<td>BS11C0</td>
<td>41.6</td>
<td>—</td>
<td>BS16C0</td>
<td>31.6</td>
<td>—</td>
</tr>
<tr>
<td>BS11(FR)C7</td>
<td>20.6</td>
<td>24.6</td>
<td>BS16(S)C3</td>
<td>39.6</td>
<td>44.8</td>
</tr>
<tr>
<td>BSTLC0</td>
<td>—</td>
<td>37.7</td>
<td>BSTL(D)C3</td>
<td>7.0</td>
<td>25.4</td>
</tr>
</tbody>
</table>

$\bar{X}$ C0 37.4 — 46.1 —  
$\bar{X}$ Ci 19.2 27.1 2.2 31.1

$^a$Estimates of inbreeding depression were calculated as $(\bar{X}_0 - \bar{X}_F) / \bar{X}_0 \times 100$, where $F = 50\%$. The $S_1$ generations of each population were obtained by self pollination of 100 plants and bulking equal quantities of seed from each self-pollinated ear for evaluation trails.


less than for BSSSC0. Relative to BSSSC0, BS13(S)C4 $S_1$ generation was 14.4% greater yielding than the non-inbred BSSSC0 population. Two sets of selection studies (half-sib and inbred) were conducted in BSKC0 (Tanner and Smith, 1987). After eight cycles of selection, ID estimates for BSK(H)C8 and BSK(S)C6 were 30.4 and 26.4%, respectively, compared with 54.9% for the original unselected population, BSKC0. Except for BSCB1 and BS16, it seems selection was effective in changing allele frequencies either by an increase in frequencies of desirable alleles for grain yield or by a decrease in frequencies of deleterious recessive alleles. BSCB1 was synthesized by intermating 12 inbred-lines with above average resistance to the European corn borer with limited attention considered for their combining ability. BS16 is a strain of ETO Composite from Colombia, South America, that was adapted to temperate environments and mass selection emphasized only earlier flowering. Allele frequencies for greater grain yield may have been less than 0.5 before selection initiated for grain yield. On the average, estimates of ID for the selected populations were 10–15% less than for the C0 populations (Table 9.8).

Table 9.9 shows the example of ID based on both $S_1$ and $S_7$ progenies for grain yield and ear length in BSSS.
### Table 9.9 Inbreeding depression for two traits in BSSS maize population

<table>
<thead>
<tr>
<th>Generation</th>
<th>$F$</th>
<th>Grain yield (q/ha)</th>
<th>Ear length (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0</td>
<td>0.000</td>
<td>45.2</td>
<td>15.9</td>
</tr>
<tr>
<td>S1</td>
<td>0.500</td>
<td>26.8</td>
<td>13.9</td>
</tr>
<tr>
<td>S2</td>
<td>0.750</td>
<td>20.0</td>
<td>13.5</td>
</tr>
<tr>
<td>S3</td>
<td>0.875</td>
<td>15.8</td>
<td>12.7</td>
</tr>
<tr>
<td>S4</td>
<td>0.938</td>
<td>15.3</td>
<td>12.9</td>
</tr>
<tr>
<td>S5</td>
<td>0.969</td>
<td>14.7</td>
<td>12.8</td>
</tr>
<tr>
<td>S6</td>
<td>0.984</td>
<td>14.1</td>
<td>12.9</td>
</tr>
<tr>
<td>S7</td>
<td>0.992</td>
<td>13.0</td>
<td>12.5</td>
</tr>
<tr>
<td>ID (S₁)</td>
<td></td>
<td>40.7</td>
<td>12.7</td>
</tr>
<tr>
<td>ID (S₇)</td>
<td></td>
<td>71.2</td>
<td>20.1</td>
</tr>
</tbody>
</table>

#### 9.6 Frequency of Useful Lines

Lindstrom (1939) summarized a survey of the maize breeders in the USDA and in 24 US Agriculture Experiment Stations to determine the number of inbred-lines that had been isolated. He reported that 27,641 lines had been selfed for 1–3 years and that only 677 (2.4%) were presumably useful lines. His assessment of the quality of the inbred-lines was rather pessimistic because every one of the useful lines was seriously defective for one or more traits, including poor yields. Lindstrom (1939) concluded that probably 100,000 lines had been tested for at least 3 years but very few were useful. As a consequence, he questioned the use of self-fertilization as the proper inbreeding method for isolating inbred-lines. Almost 70 years later pedigree selection of elite × elite combinations within heterotic groups continues to be the most popular methodology for inbred-line development and it is unlikely that the relative number of useful lines would have increased with larger phenotyping and genotyping capabilities. One of the questions today is how many unique lines have increased the genetic diversity of the commercial hybrids available across maturity regions. Even though line recycling and sample sizes are indirectly affected by inbreeding this is a germplasm question for Chapter 11.

Public institutions were the principal developers of maize inbred-lines and industry relied heavily on them until the 1980s. Since World War II and particularly during the 1960s and 1970s there has been a tremendous expansion in the commercial sector in number and size of breeding programs primarily designed to isolate and test new inbred-lines followed by the isolation of transgenes for what originally was known as genetically modified organisms (GMOs) in the 1990s (e.g., Bt, RR). These hybrids had a technology fee associated to them for each single-gene trait delivered. Hybrids with eight events combined will be available in 2010 but still are based on pest and herbicide resistance only. As a consequence, the inbred–hybrid concept became a bit more complex where major tasks are involved such as conversions, testing, integration of conversions and line development, data management,
The cooperative state-federal maize breeding program at Ames, Iowa, included on the average about 360 testcrosses each year, which usually involved the initial testing of $S_2$ or $S_3$ lines for combining ability for yield, lodging, and grain moisture at harvest. These lines have survived previous selection for various agronomic traits in the breeding nursery and in some instances have had some previous testing in recurrent selection programs. The state maize breeding program at Fargo, North Dakota, currently includes 960 new testcrosses each year at the $S_0$ level (in addition to the re-testing of the top ones across breeding stages, see Chapter 1) where in-kind industry and producer cooperation is essential. If we consider 360 testcrosses as an average for each year of the 40 years between 1939 and 1979, 14,400 lines had merited testing in hybrid combination. If we extrapolate from the program to include 25 public and 25 private breeding programs that were available at that time before consolidation in both the public and private sectors, we have an average of 18,000 lines tested each year or 720,000 since 1939. The numbers seem large, but they may be conservative because several of the larger companies have several breeding stations to select lines for use in particular environments. The number of lines that have been inbred, selected, and advanced to testcrossing may be in the neighborhood of 1 million, most of them related genetically though. Surveys reported by Sprague (1971), Zuber (1975), and Darrah and Zuber (1986) list the inbred-lines and the extent of their use in producing commercial hybrid seed. Due to intellectual property it is currently challenging to track down public lines as they become coded in Foundation Seed Companies or they are sold by Research Foundations representing ownership in public Institutions. At that time, the inbred-lines developed by public agencies included 38 that were used in 0.1% or more of the seed requirements for the 1976 US crop year (Zuber, 1975). The relative proportion of the 38 useful inbred-lines relative to the total tested by public agencies is only about 0.01%, which means that only 1 in 10,000 $S_2$ or $S_3$ lines tested were eventually used to any extent in commercial hybrids. Although maize breeders, both public and private, have been very active in developing and testing inbred-lines, the frequency of developing new and unique inbred lines of commercial value is indeed very low. Smith (1988), Mikel and Dudley (2006), and Mikel (2008) have studied the diversity and genetic origins of US maize inbred-lines. Germplasm of a few lines, such as B14, B37, B73, C103, Idt, and Oh43, persists decades after they were originally developed because of recycling of elite lines via pedigree selection (see Chapter 1). The method of linking germplasm improvement (e.g., recurrent selection) with inbred-line development followed by recycling has been very effective (Duvick et al., 2004). There are several pitfalls in calculating the estimates of useful lines though. In some ways the estimate of 0.01% seems too small. Again using estimates for the Iowa location in the central US Corn Belt, four inbred-lines (B37, B73, B14A, and B57) were included in the survey reported by Zuber (1975). Relative to the estimated 14,400 testcrosses for the Iowa location, the percentage of useful lines is 0.027, an estimate about three times greater than for all stations. This estimate seems extremely small, but it seems realistic compared that are eventually used to any extent in commercial
hybrids. The base of reference may not seem valid, but several of the lines (B14, B37, Oh43, and C103) were developed during this period of 40 years so that the total span of time seems correct. Also several of the 38 lines included in the survey are recoveries (A632, A619, Va26, H84, B68, etc.) of lines (e.g., recycling) and may inflate the actual number of usable lines developed. Weighing the advantages and disadvantages for calculating the percentage of useable lines from inbreeding and testcrossing programs, 0.01% may be a good estimate, but it is considerably lower than the 2.4% reported by Lindstrom (1939). This number could increase if public breeders concentrate on niche low-investment markets not served by industry, usually affected by significant environmental challenges (e.g., very short seasons, frost, salt and drought stresses).

9.7 Types of Hybrids Produced from Inbred-Lines

Shull’s (1908, 1909, 1910) original concept was the production and growing of single-cross hybrids, but the costs of seed production seemed to limit its usefulness. This limitation was overcome with Jones’s (1918) suggestion that double-cross hybrids can be produced from two single-cross hybrids to reduce the costs of seed production. Consequently, double-cross hybrids became rapidly accepted in the USA (see Chapter 1). Single-cross hybrids, however, gradually replaced double-cross hybrids in the US Corn Belt and other areas of the world since 1960. Practically 100% of the US Corn Belt is currently planted to single-cross hybrids. In addition three-way crosses and modified single crosses have been used in some instances, mainly because some problems of seed production are alleviated by use of the single cross as a seed parent. This procedure has also been useful for early-generation hybrid testing in maize breeding programs of the northern USA and Canada. In the modified single cross either one or both parents are crosses of related lines, i.e., parent lines have a common parent in their ancestry. Usually the seed parent is produced from a cross of two closely related lines to enhance the quantity of seed produced because of some hybrid vigor from the cross. The degree of ancestry of the parents included in a modified single cross is not exact and varies among modified single crosses. For example, one may consider (H84 × H93)Va26 or (LH176 × LH177)ND2000 modified single crosses. Both H84[(B37 × GE440)HtHt] and H93 [(B37 × GE440)B374Ht− Ht] have B37 in their ancestry in different proportions, some hybrid vigor is expressed in the cross of H84 × H93, and both H84 and H93 combine well with Va26. Also, both LH176 and LH177 (known as sister lines due to their close relatedness) are ‘Holdens’ lines that have LH82 in their background. LH82 is an older line whose PVP protection has expired. Every year, several industry lines become available after ~20 years of protection. Even though they seem old for a breeder potential new successful hybrid combinations are possible.

Different types of maize hybrids have advantages and disadvantages. Initially, the low vigor and productiveness of the inbred-lines available discouraged the production of single-cross hybrid seed at an acceptable cost. Recycling of lines and selection have produced lines that are more productive than the earlier derived
lines, but increased usage of fertilizers and herbicides and improved husbandry practices also have contributed to the feasibility of producing single-cross hybrid seed (Duvick, 1999; Duvick et al., 2004). Schnell (1975) discussed three main aspects in his comparisons of different types of hybrids:

1. Uniformity,
2. Yield, and

A fourth aspect was added:
4. The relative simplicity of selecting and testing the three types of hybrids.

Simplicity was definitively essential for the use of single crosses.

Uniformity has been an important factor for the acceptance of single crosses. If good practices are used in the production of a single-cross hybrid, single crosses are uniform genotypically and phenotypically. A field of a single-cross hybrid is very attractive to the producer because of its uniformity of appearance, maturity, and harvesting characteristics. A single cross is genetically homogeneous, which may be a disadvantage. If large areas are planted to the same single-cross hybrid, uniformity will be evident not only within a given field but also in the large area. This situation, for a particular maturity zone, existed when the 1970 *Bipolaris maydis* outbreak occurred in the USA. It was caused by the use of T-cytoplasm in the production of the hybrids. It is difficult to fathom that a situation similar to the use of T-cytoplasm would occur over as an extensive area for genotypic uniformity. The extreme uniformity of single crosses, however, does cause some concern about the host–pathogen relationship within each field and area that could develop as genetic diversity among and within single crosses is reduced. Planting different hybrids that are genetically distinct would solve part of this problem.

One of the advantages often given for the use of single-cross hybrids is that greater yields are possible. Intuitively, this seems correct because it is easier to identify two inbred-lines that are superior yielding in crosses than to identify three or four in other types of hybrids. Superior performing single crosses are identified in the first stages of testing. Single-cross data were used to predict the performance of three-way and double-cross hybrids, usually based on the average performance of the non-parental single crosses (Chapter 10). A summary of comparisons among types of hybrids given in Table 9.10 includes number of hybrids included in comparisons, average yield, and relative yields, with single-cross hybrids the base of comparison. There is some variation in advantage of one type of hybrid relative to another, but the differences generally are not large. Weighted and unweighted averages were very similar in Table 9.10.

Average comparisons, however, can be misleading in some instances, and two examples illustrate some of the comparisons. Schnell (1975) examined in some detail the data reported by Weatherspoon (1970), who used nine unrelated inbred-lines to produce the 36 possible single crosses and balanced sets of 36 three-way and double crosses. Schnell’s summary of Weatherspoon’s data (Table 9.11) shows that the single crosses had greater average yield, greater standard deviation, and greater
Table 9.10 Summary of the average and relative yields (q/ha) for single, three-way, and double-cross hybrids of maize

<table>
<thead>
<tr>
<th>Source</th>
<th>Single crosses</th>
<th>Three-way crosses</th>
<th>Double crosses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yield No. %</td>
<td>Yield No. %</td>
<td>Yield No. %</td>
</tr>
<tr>
<td>Doxtator and Johnson (1936)</td>
<td>47.0 6 100</td>
<td>51.4 2 109.4</td>
<td>44.9 3 95.5</td>
</tr>
<tr>
<td>Stringfield (1950)</td>
<td>53.1 6 100</td>
<td>53.1 12 100.0</td>
<td>51.2 3 96.4</td>
</tr>
<tr>
<td>Jones (1958)</td>
<td>44.4 317 100</td>
<td>— — —</td>
<td>45.0 483 101.4</td>
</tr>
<tr>
<td>Jugenheimer (1958)</td>
<td>65.4 6 100</td>
<td>62.3 12 95.3</td>
<td>61.3 3 93.8</td>
</tr>
<tr>
<td>Sprague et al. (1962)</td>
<td>68.1 60 100</td>
<td>66.6 60 97.8</td>
<td>— — —</td>
</tr>
<tr>
<td>Sprague and Thomas (1967)</td>
<td>75.3 15 100</td>
<td>73.8 60 98.0</td>
<td>— — —</td>
</tr>
<tr>
<td>Eberhart and Hallauer (1968)</td>
<td>71.0 6 100</td>
<td>70.7 12 99.6</td>
<td>70.4 3 99.2</td>
</tr>
<tr>
<td>Eberhart and Russell (1969)</td>
<td>71.0 45 100</td>
<td>— — —</td>
<td>69.6 45 98.0</td>
</tr>
<tr>
<td>Weatherspoon (1970)</td>
<td>65.1 36 100</td>
<td>62.0 36 95.2</td>
<td>60.3 36 92.6</td>
</tr>
<tr>
<td>Wright et al. (1971)</td>
<td>47.8 150 100</td>
<td>48.4 600 101.2</td>
<td>— — —</td>
</tr>
<tr>
<td>Stuber et al. (1973) - U</td>
<td>59.4 84 100</td>
<td>60.4 168 101.7</td>
<td>59.2 42 99.7</td>
</tr>
<tr>
<td>S</td>
<td>60.6 84 100</td>
<td>61.4 168 101.3</td>
<td>61.1 42 100.8</td>
</tr>
<tr>
<td>Lopez-Perez (1977) - S</td>
<td>92.2 — 100</td>
<td>87.4 — 94.8</td>
<td>89.9 — 97.5</td>
</tr>
<tr>
<td>FS</td>
<td>89.0 — 100</td>
<td>89.3 — 100.3</td>
<td>88.0 — 98.9</td>
</tr>
<tr>
<td>Unweighted</td>
<td>65.3 15 100</td>
<td>66.0 13 100.4</td>
<td>64.3 12 98.6</td>
</tr>
<tr>
<td>Weighted</td>
<td>53.5 821 100</td>
<td>55.5 114 103.8</td>
<td>49.8 663 93.1</td>
</tr>
<tr>
<td></td>
<td>64.0c 354 100</td>
<td>63.3c 542 98.9</td>
<td>62.5c 180 97.7</td>
</tr>
</tbody>
</table>

*a*Hybrids produced among selected and unselected lines

*b*Hybrid bulks from lines developed by selling (S) and full-sibbing (FS) systems of inbreeding

*c*Averages omitting Jones (1958) and Wright et al. (1971) for single crosses; Wright et al. (1971) for three-way crosses, and Jones (1958) for double crosses

range from low to high crosses than did double crosses, with three-way crosses intermediate. The best single cross was 13.8 and 8.6 q/ha greater than the best double and three-way crosses, respectively. Expected greatest yield also was obtained for each type of cross predicted on the basis of the largest deviate expected in a sample of 36 crosses. In all instances, it seems that the best single crosses have a striking advantage over the best three-way and double crosses. In predicting the best cross for each type, a constant $k$ value was used. Schnell emphasized, however, that all possible crosses among the nine inbred-lines were produced and tested but only a balanced set of the possible three-way and double crosses were included.
9.7 Types of Hybrids Produced from Inbred-Lines

Table 9.11 Yield (q/ha) distributions of balanced sets of 36 single, three-way, and double crosses (Schnell, 1975)

<table>
<thead>
<tr>
<th>Crosses</th>
<th>Average</th>
<th>Standard deviation</th>
<th>Extremes</th>
<th>Expected greatest cross$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Least</td>
<td>Greatest</td>
</tr>
<tr>
<td>Single</td>
<td>65.1</td>
<td>8.8</td>
<td>43.6</td>
<td>81.5</td>
</tr>
<tr>
<td>Three-way</td>
<td>62.0</td>
<td>6.2</td>
<td>47.7</td>
<td>72.9</td>
</tr>
<tr>
<td>Double</td>
<td>60.3</td>
<td>3.8</td>
<td>54.0</td>
<td>67.7</td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three-way</td>
<td>65.1</td>
<td>6.4</td>
<td>47.4</td>
<td>80.0</td>
</tr>
<tr>
<td>Double</td>
<td>65.1</td>
<td>4.8</td>
<td>52.5</td>
<td>79.1</td>
</tr>
</tbody>
</table>

$^a$Expected cross based on: $\bar{X} + k\hat{\sigma}_x$, where $k = 2.12$, the greatest expected deviate in a sample of 36

The single-cross information was used to predict the 252 possible three-way and 378 double crosses. Greatest yields of the predicted three-way and double crosses (Table 9.11) are only 1.5 and 2.4 q/ha, respectively, less than the single crosses. Predictions may be biased with regard to the average and individual crosses, but Otsuka et al. (1972) indicated that yields of the better three-way and double crosses were underestimated from the non-parental single crosses.

Jones (1958) made a comparison among 317 single and 483 double crosses available from field trials grown in Iowa in 1951 (Table 9.12). Single crosses were made by crossing in all combinations a series of highly selected inbred-lines, and double crosses were produced from a series of single crosses. Mean yields of the two types of hybrids were not different, but there were differences in the range and distribution of the two types. No double crosses were represented in the two extreme low and high classes, and single crosses had a distinctly bimodal distribution. Single crosses also had a greater standard deviation than double crosses. Although only two types of hybrids were included, comparisons for single and double crosses were similar to those reported by Weatherspoon (1970). Therefore, single crosses had a greater variation in yield and a greater standard deviation.

The differences in yield among the three types of crosses do not seem to be as great as expected. Cockerham (1961) showed an advantage in selecting among single crosses compared with three-way and double-crosses for lines developed from a specified population regardless of the type of gene action expressed in the crosses. Stuber et al. (1973) presented data that agreed with the genetic theory presented by Cockerham (1961). If only additive genetic effects are important, the

Table 9.12 Frequency distributions for 317 single and 483 double crosses for yield (Jones, 1958)

<table>
<thead>
<tr>
<th>Class centers (q/ha)</th>
<th>Single</th>
<th>Double</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23.8</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>26.9</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>30.0</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>33.1</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>36.2</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>39.4</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>42.5</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>45.6</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>48.8</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>51.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>55.0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>58.1</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>61.2</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>64.4</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>$\bar{X}$</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
expected selection advance is twice as great among single crosses as among double crosses. As the relative importance of non-additive effects (dominance and epistasis) increases, the relative advantages of selecting among single crosses are even greater.

It is shown in Chapter 5 that the estimation of epistatic variance for maize populations is not satisfying. However, Dudley and Johnson (2009) show prediction power is increased with the inclusion of epistasis. Comparisons among means of genetic populations formed from specific pairs of genotypes invariably show evidence for significant net epistatic effects. If epistatic effects contribute to the heterosis expressed in a cross of two inbred-lines, single crosses are superior to three-way and double crosses because the unique epistatic combinations of the single cross are disrupted in the gametes of the single-cross parents used to produce the three-way and double crosses. If unique combinations of epistatic effects are important in single crosses, a greater range of yields is expected for single crosses than for three-way and double crosses, as demonstrated by Jones (1958) and Weatherspoon (1970).

Sprague et al. (1962) and Sprague and Thomas (1967) made comparisons of single and three-way crosses for two sets of inbred-lines. Sprague et al. (1962) used a highly selected set of six inbred-lines from different sources and found that 18 of 20 single-cross sets yielded more than the corresponding three-way sets. The highly selected lines were identified on the basis of their performance in crosses. Selection and testing apparently identified lines that contributed unique combinations of epistatic effects to their crosses. Sprague and Thomas (1967) used a set of unselected inbred-lines developed from the Midland variety and found 11 of 19 instances where the single crosses were higher yielding than the three-way crosses. The approximate equivalence of the single- and three-way cross yields produced from their unselected lines agrees with the expectation that any selection involved was neutral with respect to types of gene action. The data of Lopez-Perez (1977) show similar yields of the three types of hybrids produced from unselected lines.

The contrasts of the results for specific studies and the averages of Table 9.11 are not as serious as they appear. It seems epistasis is important in single crosses, but sampling influences results. For a fixed set of lines, the number of possible combinations of three-way and double crosses is much greater than for single crosses. If all possible three-way and double crosses were tested extensively, it may be possible to identify three-way and double crosses that also have unique combinations of epistatic effects. The unique combinations may be different from those expressed in single crosses because of genetic recombination of single-cross parents. Because the number of possible crosses (e.g., for nine lines, we have 36 single-, 252 three-way, and 378 double crosses) increases dramatically for a fixed set of inbred-lines, prediction methods of Jenkins (1934) have been used to identify the particular three-way and double crosses for testing. Otsuka et al. (1972) and Stuber et al. (1973) found that use of non-parental single crosses underestimated three-way and double-cross performances, but they concluded that the underestimates seem small and the biases due to epistasis and genotype by environment interactions were similar. The same conclusions were reported by Eberhart and Hallauer (1968). None of the studies
recommended use of more complex procedures to predict three-way and double 
crosses.

Single crosses are heterozygous, but they are homogeneous with respect to the 
genotype because each plant within the hybrid is genetically the same. The lack of 
genetic variability within a single cross has concerned maize breeders and produc-
ers on both an individual field and an area basis. Although single crosses may be 
superior to three-way and double crosses, the consistency of performance over envi-
ronments (or stability) was also of concern. It seemed that external environmental 
factors (weather, soil, and pests) would have a greater effect on the genotypically 
uniform single crosses than on the genotypically variable three-way and dou-
ble crosses. Jones (1958) examined the comparison of single and double crosses 
and concluded that the greater genetic homeostasis of double crosses was more 
important than the problems associated with producing single crosses. Federer and 
Sprague (1947) and Sprague and Federer (1951) estimated the genotype by environ-
ment component of variance for three types of hybrids and found that the interaction 
component was greater for single crosses than for double crosses, indicating a more 
sporadic performance of single crosses in different environments and, thus, less 
stability. The mixture of related genotypes of double crosses, therefore, showed 
smaller hybrid by environment interactions than the single crosses and presumably 
had greater stability over environments than single crosses. Since no comparisons 
were made for individual hybrids within types, Gama and Hallauer (1980) evaluated 
single-crosses and found no differences in relative stability of single-cross hybrids 
produced from selected and unselected inbred-lines.

Eberhart and Russell (1969) used their stability analysis (Eberhart and Russell, 
1966) for a diallel set of 45 single crosses and a corresponding balanced set of 
double crosses derived from 10 selected inbred-lines. The 90 hybrids were grown 
at 21 US Corn Belt environments and the traditional analysis of variance showed 
Nearly three times greater variation among single crosses (490.0) than among dou-
ble crosses (170.2). The hybrid by environment interaction was slightly greater for 
the single crosses (55.5 vs. 37.6). The stability analysis identified two single crosses 
as being as stable as any of the double crosses. These two single crosses yielded 11% 
more than four commercial single crosses and performed better than three commer-
cial double crosses (+ 13%). Comparisons of the highest yielding single and double 
crosses for the 21 environments showed that the differences among the stability 
parameters were not different. On the average double-cross hybrids were slightly 
more stable with respect to stability parameters, but single crosses were slightly 
higher yielding in all types of environments (poor, average, and good). None of 
the differences, however, was significant. Eberhart and Russell (1969) concluded 
that single crosses are as stable as double crosses over the environments tested. 
However, extensive testing over a wider range of environments would be needed 
to identify good yielding single crosses for commercial use. It is doubtful that the 
resources needed to identify stable single crosses would be any greater than needed 
for three-way and double crosses. To provide valid data for predicting three-way and 
double crosses, it was necessary to have adequate yield testing of single crosses. 
Single-cross data were used to predict the combinations of three-way and double
crosses, which were then needed to determine the best performing and most stable three-way and double crosses. The two-stage testing required for three-way and double crosses fulfilled the wider testing requirement needed to identify stable single crosses. Hence additional resources were not required to identify superior, stable single crosses.

All aspects of comparisons among single, three-way, and double crosses do not show any striking advantages for a type except for the simplicity of testing and identifying single crosses. Good yielding single crosses with desired agronomic traits are usually easier to identify than the more complex hybrids, but this does not imply that equally good yielding three-way and double crosses could not be identified after extensive testing. As shown in Table 9.11, good yielding three-way and double crosses can be identified if sampling of the possible three-way and double crosses is included. This has been prohibitive in most instances for a given set of lines because the number of possible combinations increases rapidly. However, modified single crosses have been a good alternative option.

The advantage of simplicity for single crosses is evident in the mechanical procedures of breeding and production. Greater breeding effort can be allotted to line development if resources are not needed to produce the three-way and double-cross seed for testing. Testing is simpler because only one-stage testing is required rather than two stage for the more complex hybrids. Production of single crosses is simpler because they require only three isolation fields (two for foundation seed stocks and one to produce hybrid seed), whereas perhaps seven isolation fields (four for the foundation seed stocks, two for parental single crosses, and one to produce double-cross hybrid seed) are needed to produce double crosses. The difference in number of isolation fields is usually not as great for the two types of hybrids because foundation seed stock organizations are available to produce the parental stocks. Production of single-cross seed is certainly more expensive per unit area than double-cross seed, but this disadvantage also must be considered with respect to maintenance of two seed stocks for single crosses vs. six for double crosses. The other advantages of single crosses and the ratio of seed costs to other production costs have encouraged the use of single-cross hybrids.

9.8 Heterozygosity and Performance

Like the relation between rates of inbreeding depression and inbreeding systems studies relate the level of heterozygosity and performance of some quantitative trait as evidence for the presence (or absence) of epistatic effects. The basis of this relation traces back to Wright’s (1922b) investigations of heterosis and inbreeding depression in guinea pigs. These experiments were the basis for Wright’s general conclusion that the change in vigor is directly proportional to the change in heterozygosity in the population. Wright summarized his findings by stating, ‘A random-bred stock derived from $n$ inbred families will have $(1/n)$th less superiority over its inbred ancestry than the first cross or a random-bred stock from
which the inbred families might have been derived without selection.’ The results of Wright can be summarized mathematically as

\[ F_2 = F_1 - \frac{(F_1 - P)}{n} \]

where \( n \) is the number of inbred parents and \( P \) is the mean of all the parents. For single-cross hybrids, \( n \) would equal two. Falconer and Mackay (1996) showed heterosis equals \( y^2 d \) for one locus. If the \( F_1 \) is self-pollinated the change on inbreeding is \(-2pqdF \) or heterosis of \( F_2 \) is \((\frac{1}{2})y^2 d \). Application of the heterozygosity–performance relation has been reported in maize. The effects of inbreeding were discussed by Gilmore (1969).

Kiesselbach (1933) generated data that he interpreted to support Wright’s conclusion. Yield data from \( F_1 \) hybrids with 2, 4, 8, and 16 inbred parents and their \( F_2 \) generations seemed to indicate that yield was highly correlated with heterozygosity. Yield data of the inbred parents were not available, but Kiesselbach (1933) summarized his findings that the decrease in yield of any hybrid in the \( F_2 \) generation is equal to half the difference in yield between the \( F_1 \) and the mean yield of the open-pollinated varieties. The reduced yields of the \( F_2 \) generations were attributed to a reduction in the number of favorable growth factors as a consequence of close breeding. Kiesselbach further concluded that for hybrids in which the parents were composed of equal numbers of parent inbreds, the reduction in yield tended to be inversely proportional to the number of lines involved.

Neal (1935) used Wright’s formula to predict the performance of advanced generations in maize hybrids. He used 10 single, 4 three-way, and 10 double-cross hybrids; their \( F_2 \) generations; and \( F_3 \) generations of six of the single crosses to test the validity of Wright’s relation in maize. Neal reported that the single, three-way, and double crosses lost 29.5, 23.4, and 15.8% vigor, respectively, in the \( F_2 \) generation relative to the \( F_1 \), which agreed with the predicted losses of 31.0, 21.0, and 15.3% for the three respective types of hybrids. Neal concluded that vigor decreased according to Wright’s formula. Wright’s formula is based on the assumption of arithmetic gene action. Powers (1941) reanalyzed Neal’s data using arithmetic and geometric models. The best agreement between observed and predicted values was obtained with the arithmetic model.

Kinman and Sprague (1945) also studied the type of model involved in loss of vigor with advanced generations. They compared the observed and expected performance of 10 maize inbred-lines, the 45 possible single crosses, and their \( F_2 \) generations by use of arithmetic and geometric models. They also concluded that the arithmetic model was a closer approximation to the observed data than was the geometric model. They emphasized, however, that the calculation of an arithmetic value does not imply that all the genetic factors involved in the expression of a particular trait operate in an additive fashion as, for instance, cancellation of inter-locus interactions may prevent their detection. The studies of Neal (1935) and Kinman and Sprague (1945) verified the hypothesis that performance was a linear function of the percent heterozygosity based on the use of inbred-lines and on the \( F_1 \) and \( F_2 \) generations corresponding to 0, 100, and 50% heterozygosity, respectively.
With these three levels of heterozygosity, there was no indication that the observed results deviated from a linear model. Subsequent studies included additional generations with levels of heterozygosity intermediate to the three levels of 0, 50, and 100%. Stringfield (1950) used combinations of four inbred-lines to produce seven genetic populations that represented four levels of heterozygosity (0, 50, 75, and 100%). He concluded that the heterozygosity–performance relation was curvilinear for maturity and ear height and less so for grain yield. The curvilinear relation indicated that as 100% heterozygosity was approached the added increments of heterozygosity had a lesser effect. He hypothesized that the added genes may be duplicating the functions of those already present, which agrees with the diminishing rate of returns hypothesis suggested by Rasmusson (1934). In a similar study, Sentz et al. (1954) included five levels of heterozygosity in two maize populations. The heterozygosity–performance relation for seven traits over four environments showed a significant deviation from linearity for all traits except maturity and plant height in one population and ear diameter in the second. A curvilinear relation was noted between the 25 and 75% levels of heterozygosity for all traits except ear number. Within an individual environment, the responses were similar in both populations for yield and maturity but the responses differed among environments, indicating interactions with environments. Sentz et al. concluded that the curvilinear response to levels of heterozygosity was evidence for epistatic gene action, but they also acknowledged that the genetic base of their materials was restricted.

Martin and Hallauer (1976) examined the relation between four levels of heterozygosity and five traits for four groups of inbred-lines. The lines in each group were assigned as follows: type I, lines from open-pollinated varieties (first cycle); type II, lines from pedigree crosses or improved populations (second cycle); type III, good lines based on good vigor and general combining ability; and type IV, poor lines based on poor vigor and combining ability. Each of the four groups of lines included seven lines, the 21 possible F\textsubscript{1}, F\textsubscript{2}, and F\textsubscript{3} generations among the seven lines, the 21 backcrosses to each of the lines, and the 21 backcrosses selfed for each of the inbred-lines. Data were collected in two replicate experiments conducted in five environments. The relation between level of heterozygosity and mean performance was determined for each environment and combined across the five environments. Sums of squares for levels of heterozygosity were partitioned into those due to linear, quadratic, and lack-of-fit components. Some instances of detectable epistatic effects were noted in all types for each trait in each environment. The frequency of detectable epistatic effects, however, was much less than that of significant linear effects. Yield, for example, combined across the four types and the five environments had 84 instances of significant ($P \leq 0.01$) linear, three quadratic, and one lack-of-fit mean squares. If the linear model assumes that individual loci contribute their effects independently of all other loci and that the quadratic and lack-of-fit mean squares indicate evidence of epistatic effects, the occurrence of net epistatic effects for yield was very low. Occurrence of significant epistatic effects was greater in the individual environments than in the combined analysis across the five environments. In addition to testing mean squares for the presence of epistatic effects, Martin and Hallauer (1976) determined the relative proportions
of levels within crosses sums of squares that were due to the linear, quadratic, and lack-of-fit sources of variation. For yield, the linear model accounted for 99.0, 97.8, 98.0, and 97.9% of the total variation for the four types of lines; the quadratic was 1.5% or less in all types. In 15 of 20 instances for all traits and for all types, the linear model accounted for 90% or more of the total variation. Figure 9.5 shows the relation of yield to level of heterozygosity, and a linear trend is evident for each of the four types of lines. Although at least one instance of significant epistatic effect was detected in each type for each trait, the relative contribution of net epistasis to the mean though seems small.

Most of the heterozygosity–performance relation studies included crosses among selected inbred-lines as their experimental material. Data also are available for crosses generated from open-pollinated varieties. Pollak et al. (1957) used three open-pollinated varieties and their F₁, F₂, and backcross populations to establish three levels of expected heterozygosity. Based on expectations, the backcrosses would equal the F₂ generations and the F₂ would be midway between the F₁ and the mean of the two parent varieties. The F₂ generations, for yield, were intermediate to the means of the parents and the F₁ for the three crosses, with the backcrosses similar to the F₂. All comparisons except one were within one standard error of the means. As a consequence, Pollak et al. concluded that the evidence does not suggest the presence of important epistatic combinations of genes conditioning yield. Robinson and Cockerham (1961) related yield and ear height to level of heterozygosity for two open-pollinated varieties, their F₁ and F₂ generations, and the selfs

![Fig. 9.5 Relation between level of heterozygosity and yield for four types of inbred-lines evaluated across five environments (Martin and Hallauer, 1976)](image-url)
of each. There was no significant deviation from their proposed additive genetic model with inclusion of dominance. Moll et al. (1965) included open-pollinated varieties and their F1 and F2 generations for different levels of genetic divergence, based on geographical origin of the parental varieties. Per plant means of the parents (159.8 g), the F1 generation (196.3 g), and the F2 generation (179.3 g) show that the F2 mean deviates only 0.7% from the average (178.0 g) of the parents and the F1 means. Hence, it seems the presence of net epistatic effects was, on average, not important for these variety crosses. For the F1 and F2 generations of the varieties included by Moll et al. (1965), the F2 averaged 8.9% lower yields than the F1 crosses. Pollak et al. (1957) found the F2 deviated from the mid-parent by only 1.2% and the F2 mean, on the average, was 6.0% lower than the F1 mean.

Shehata and Dhawan (1975) determined the inbreeding depression for three sets of diallel crosses, each set including 10 parents that were either open-pollinated or synthetic varieties. Average inbreeding for each of the three diallel series of crosses was 10.2, 11.4, and 11.2%, comparing the F2 with the F1 generations. The deviations of the F2 generations from the mid-parents of the parental varieties and their F1 variety crosses were small in all instances and were interpreted as evidence that net epistatic effect in yield was small relative to the additive effects among loci. The decreases from inbreeding F1 generations to produce F2 generations generally were less for the varieties than for specific crosses of a pair of inbred-lines (Table 9.13). Although the data are limited, the frequency of significant epistatic effects was greater for crosses of inbred-lines than for crosses among varieties. This difference seems reasonable because of the nature of the material included in the crosses. For varieties this would involve an array of genotypes included in the crosses, whereas crosses of inbred-lines would involve the expression for two specific genotypes. Martin and Hallauer (1976) found differences in frequencies of detectable epistasis among four types of inbred-lines based on origin of the lines included in the four types. The effects of inbreeding depression also seem greater for F1 generations of inbred lines than for variety crosses. The results from studies of F1 and F2 generations produced from inbred-lines for two different eras show a large reduction from the F1 to the F2 generations. Kiesselbach (1922) reported the F2 generations were 46.7% lower yielding than the F1 generations. Martin and Hallauer (1976), for 84 crosses among 28 lines, found the F2 generations were 30.1% lower yielding than the F1 generations. Some of the differences can be attributed to the improvements made in the yield capacity of the inbred-lines used in the crosses. Kiesselbach presented data for 1 year and found the F1 hybrids were 404.7% higher yielding than the inbred-lines, while Martin and Hallauer (1976) showed the hybrids were only 184.2% higher yielding than the inbred-lines. For an unselected group of lines developed from Iowa Stiff Stalk Synthetic, Lopez-Perez (1977) found that the F2 generations yielded 35.5% less than the F1 generation single crosses.

Data estimating the rates of inbreeding depression and comparing the means of different generations with level of heterozygosity suggest net epistatic effects are small relative to additive effects among loci regardless of the methods and genetic materials used in the comparisons. The greatest proportion of the total variation usually was accounted for by the linear model, which, particularly for yield, was
<table>
<thead>
<tr>
<th>Source</th>
<th>Type of cross</th>
<th>Parents (P)</th>
<th>F₁</th>
<th>(F₁−P)/P×100</th>
<th>F₂</th>
<th>(F₁−F₂)/F₁×100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inbred Parents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiesselbach (1922)</td>
<td>SC</td>
<td>6.4</td>
<td>32.3</td>
<td>404.7</td>
<td>15.7</td>
<td>51.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>32.2</td>
<td>—</td>
<td>18.4</td>
<td>42.8</td>
</tr>
<tr>
<td>Kiesselbach (1930)</td>
<td>SC</td>
<td>15.0</td>
<td>35.6</td>
<td>137.3</td>
<td>24.0</td>
<td>32.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>—</td>
<td>33.4</td>
<td>—</td>
<td>22.6</td>
<td>32.3</td>
</tr>
<tr>
<td>Richey et al. (1934)</td>
<td>DC</td>
<td>—</td>
<td>49.1</td>
<td>—</td>
<td>41.5</td>
<td>15.5</td>
</tr>
<tr>
<td>Neal (1935)</td>
<td>SC</td>
<td>14.8</td>
<td>39.2</td>
<td>164.9</td>
<td>27.6</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 W</td>
<td>40.1</td>
<td>169.1</td>
<td>30.8</td>
<td>23.2</td>
</tr>
<tr>
<td></td>
<td>DC</td>
<td>15.6</td>
<td>40.1</td>
<td>157.0</td>
<td>33.8</td>
<td>15.7</td>
</tr>
<tr>
<td>Kinman and Sprague (1945)</td>
<td>SC</td>
<td>17.6</td>
<td>49.9</td>
<td>183.5</td>
<td>31.8</td>
<td>36.3</td>
</tr>
<tr>
<td>Stringfield (1950)</td>
<td>SC</td>
<td>25.9</td>
<td>53.1</td>
<td>104.9</td>
<td>38.9</td>
<td>36.6</td>
</tr>
<tr>
<td>Martin and Hallauer (1976)</td>
<td>SC</td>
<td>21.5</td>
<td>61.1</td>
<td>184.2</td>
<td>42.7</td>
<td>30.1</td>
</tr>
<tr>
<td>Lopez-Perez (1977)</td>
<td>SC</td>
<td>—</td>
<td>90.6</td>
<td>—</td>
<td>58.4</td>
<td>35.5</td>
</tr>
<tr>
<td></td>
<td>3 W</td>
<td>—</td>
<td>87.4</td>
<td>—</td>
<td>55.9</td>
<td>36.0</td>
</tr>
<tr>
<td></td>
<td>DC</td>
<td>—</td>
<td>89.9</td>
<td>—</td>
<td>56.0</td>
<td>37.7</td>
</tr>
<tr>
<td><strong>Variety Parents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pollak et al. (1957)b</td>
<td>VC</td>
<td>174.0</td>
<td>202.7</td>
<td>16.5</td>
<td>190.6</td>
<td>6.0</td>
</tr>
<tr>
<td>Robinson and Cockerham (1961)b</td>
<td>VC</td>
<td>236.0</td>
<td>267.8</td>
<td>13.5</td>
<td>181.6</td>
<td>32.2</td>
</tr>
<tr>
<td>Moll et al. (1965)b</td>
<td>VC</td>
<td>159.8</td>
<td>196.3</td>
<td>22.8</td>
<td>179.3</td>
<td>8.9</td>
</tr>
<tr>
<td>Shehata and Dhawan (1975)</td>
<td>VC</td>
<td>82.2</td>
<td>101.8</td>
<td>23.8</td>
<td>91.4</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>75.4</td>
<td>89.0</td>
<td>18.0</td>
<td>78.1</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>VC</td>
<td>81.9</td>
<td>93.1</td>
<td>13.7</td>
<td>83.5</td>
<td>11.2</td>
</tr>
</tbody>
</table>

*aCrosses are designated as SC, 3 W, DC, and VC for single crosses, three-way crosses, double crosses, and variety crosses, respectively

*bGrams per plant; all other yields are expressed as quintals per hectare
greater than 90% and usually greater than 98%. As for the experiments with generation means (Chapter 5), epistatic effects are detected in special crosses but their relative frequency is low and the total variation attributed to epistasis is small. It should be reemphasized that only net epistasis is detected, which implies not that epistasis is absent (which is unrealistic) but that it must be of a cancelling nature from the comparisons of generations. Therefore, genetic models based on additivity of locus effects seem to be an adequate mathematical description of the gene action operative in the inbreeding effects in maize.

The effects of inbreeding are known, their rates for different traits determined, and various systems of inbreeding discussed. The present stage of hybrid maize in the US Corn Belt requires use of single hybrids produced from inbred-lines. Although the effects of inbreeding on phenotype are usually undesirable, use of inbred-lines to produce the high yielding, uniform hybrids will continue to ensure extensive use of inbreeding in the future. Self-fertilization for the most part and the recent use of doubled haploids seem to continue. Kiesselbach (1922) stated that ‘no elemental strain of corn has yet been found . . . , which is as vigorous or productive as the original variety from which it was derived by repeated self-fertilization.’ This statement is equally true now even though inbred-lines are much more productive than what they used to be. Improvements have been made in general productiveness and vigor of inbred-lines since the suggestion of Shull (1908). Recurrent selection procedures are effective for improving the general level of breeding populations, but inbreeding reduces the complex genotypes of the populations to their homozygous form and, consequently, vigor and productiveness will be reduced. As we continue to upgrade the general performance of breeding populations, we can expect to obtain greater yielding and more vigorous inbred-lines, but inbreeding depression will continue to be evident when self-fertilization is imposed on the complex array of genotypes within a population. Duvick (1999) and Duvick et al. (2004) have reported data for the eras from 1930 to 2000. They found that inbred-lines and hybrid yields have increased over time with the inbred yields increasing at a greater rate than hybrid yields. Consequently, the rate of heterosis expressed seemed to be increasing at a lesser rate because of the increased gains made in inbred-lines. The greater inbred-line yields can be attributed to the success breeders have had by selecting within elite line crosses to develop recycled ones that are more vigorous, have improved resistance and/or tolerance to important pests, greater drought tolerance, and greater more stable yields. There have been reports that inbred yields have exceeded 5–6 t ha$^{-1}$ (Hallauer and Carena, 2009). Recycling of elite inbred-lines is a very important phase of current maize breeding (Mikel, 2008). Nearly 100% of the current maize most productive growing regions of the world’s temperate areas are planted to single-cross hybrids. Because of economic and environmental conditions, other types of hybrids (modified single crosses, F$_2$ generation of hybrids, three-way and double-cross hybrids) and local, adapted varieties improved by selection and their hybrids have been suggested and used (Morris, 1988; Dowswell et al., 1996; Carena, 2005). Because of the expression of heterosis (Chapter 10) observed among crosses of pure line genotypes, self-fertilization is an important feature of nearly all maize breeding programs. Reducing the complex genotype of a breeding population
to its pure line (inbred-lines) components permits the identification and reproduction of genotypes that can be used to produce specific elite hybrids. This is the primary objective of inbreeding in maize.

References

References

Harvey, P. H., and J. A. Rigney. 1947. Inbreeding studies with prolific corn varieties. Department of Agronomy, North Carolina State University, Raleigh, NC.


Lopez-Perez, E. 1977. Comparisons among maize hybrids made from unselected lines developed by selfing and full-sibbing. Master’s thesis, Iowa State University, Ames, IA.


Heterosis is a phenomenon not well understood but has been exploited extensively in breeding and commercially. Hybrid cultivars are used for commercial production in crops in which heterosis expression is important. The commercial use of hybrids is restricted to those crops in which the amount of heterosis is sufficient to justify the extra cost required to produce hybrid seed.

Heterosis is the hybrid vigor manifested in hybrids and represents the superiority in performance of hybrid individuals compared with their parents. Hybrid vigor in maize is manifested in the offspring of inbred lines with high specific combining ability (SCA). Heterosis was first applied by the purposed hybridization of complex hybrid mixtures made by farmers in the 1800s (Enfield, 1866; Leaming, 1883; Waldron, 1924; Anderson and Brown, 1952). However, public scientists E. M. East and G. H. Shull developed the concept of hybrid vigor or heterosis in maize independently in the early 1900s (Shull, 1952; Wallace and Brown, 1956; Hayes, 1963). It was realized that genetic divergence of parental crosses was important for hybrid vigor expression (Collins, 1910). But, the range of genetic divergence limited the expression of heterosis (Moll et al., 1965). Heterosis can be inferred from heterotic patterns (Hallauer and Carena, 2009). A heterotic pattern is the cross between known genotypes that expresses a high level of heterosis (Carena and Hallauer, 2001). For example, it is known that inbred lines derived from Iowa Stiff Stalk Synthetic (this is a heterotic group) cross best with inbred lines derived from Lancaster Sure Crop (this is another heterotic group). Heterotic patterns became established by relating the heterosis of crosses with the origin of the parents included in the crosses (Hallauer et al., 1988). This was a consequence of diallel crosses studies on performance based on pedigree relationships. The data suggested that hybrids of lines from different germplasm sources had greater yields than hybrids of lines from similar sources. More than 50 years were needed to identify hybrid combinations that provided the highest yielding corn hybrids. Predicting the best hybrid combination is a breeding process that needs good germplasm knowledge and extensive testing. Modern research approaches based on biochemical assays (Smith et al., 1985a,b) or DNA marker data (Dudley, 1993; Stuber, 1994; Labate et al., 2004).
et al., 1997; Melchinger, 1999) have been very useful to assess genetic diversity and genetic divergence but of limited usefulness for predicting good heterotic combinations. These studies were not successful because other population properties, such as the importance of dominance genetic effects (Falconer and Mackay, 1996) or consistent validated linkages between DNA markers and quantitative trait loci (QTL) for performance (Dudley, 1993), needed to be identified. Therefore, evaluating the performance of crosses among groups based upon genetically diverse parents is essential to identify promising heterotic patterns (Melchinger, 1999). Once the heterotic patterns are established, the identification of best performing hybrids is cost-efficient. Therefore, the establishment of heterotic patterns among maize varieties has had important implications for selecting inbred lines as potential seed stocks in hybrids.

Melchinger (1999) concluded that evaluating the performance of crosses among groups based upon genetically diverse parents is essential to identify promising heterotic patterns. Due to the complexity encountered in multi-trait and multi-stage selection for economically important traits the inconsistency between genotypic and phenotypic (e.g., testcross, diallel) data shows that extensive testing of phenotypic data should be priority of maize breeding programs (Barata and Carena, 2006).

The control of parents expressing heterosis in hybrid combinations is very important. As hybrid seed products need to be purchased annually it is a business motivation. Changes in breeding methods also caused changes in seed multiplication. To take advantage of the heterosis expressed in the first-generation cross of inbred lines, new seed supplies are required for each growing season. Thus farmers are unable to use seed from their harvested crop for the following growing season. Farmers do not usually have the genetic materials, know-how, and equipment to produce hybrid seed. Production of hybrid seed, consequently, has been reserved for the specialists.

Heterosis or hybrid vigor and inbreeding depression are complementary, and the two phenomena often are observed in the same studies. Maize breeding methods have been developed to take advantage of the manifestation of heterosis in crosses of inbred lines.

Publicly financed organizations in the United States initially developed inbred lines, produced and tested hybrids, and recommended their use to farmers. Because of the rapid acceptance of hybrid maize (see Chapter 1), it soon became obvious that participation of other organizations was necessary to ensure an adequate supply of high-quality seed for each growing season. The commercial aspects of hybrid maize became very attractive because parental lines used for production of hybrids could be controlled, and commercial companies were developed that specialized in the production and sale of hybrid seed. It has become a highly competitive business and commercial companies have expanded their efforts to provide better service. They have breeding programs for developing and modifying parental inbred lines, conduct research on seed quality and seedling vigor, conduct extensive yield trials to determine the best hybrids for each area of adaptation, and have agronomists in the field to assist farmers with problems of growing hybrid maize.
The development and growing of hybrid maize can justifiably be called one of the breakthroughs and greatest accomplishments of plant breeding and agriculture. The success of hybrid maize has established an ideal that is being used by plant breeders and commercial companies for other crop species. Hybrid maize can even successfully be utilized for low-cost alternatives (e.g., population-hybrid concept, see Carena, 2005, Carena and Wicks III, 2006). In these cases the use of population hybrids does not justify the additional time and effort required to extract easily reproduced lines for hybrid development (Darrah and Penny, 1975).

Zirkle (1952) and Goldman (1999) summarized early work on the occurrence of heterosis of maize, as well as of other plant species. Köelreuter (1766), Knight (1799), and Gärtner (1849, p. 791) investigated and described plant hybridization, but apparently the first correct interpretation of hybridization in maize was provided by a letter written by Cotton Mather in 1716. Darwin (1877, p. 482) was the first to conduct experiments comparing selfed and crossed plants from the same stocks of maize; he found that plant height of crossed plants was greater than selfed plants at the juvenile (19%) and mature (9%) plant stages.

Beal (1880) was aware of Darwin’s results and conducted an experiment that is analogous to present-day methods used for maize hybridization on a large-scale basis. Beal’s parental stocks, however, were open-pollinated varieties rather than the inbred lines commonly used today. Beal collected two stocks of maize that were similar but had been grown about 160 km apart for an unknown number of years. The two stocks were grown in the same field; one was detasseled and the other was used as the male parent. The hybrid seed harvested from the detasseled rows was planted and was found to be 51% superior in yield to the original varieties. Beal suggested that use of this procedure was one method to increase yields of maize, but later developments overshadowed widespread usage of variety hybrids in the United States. Beal’s results were confirmed by Sanborn (1890), McClure (1892), Morrow and Gardner (1893), and Webber (1900, 1901).

Because of the superiority of variety hybrids, it has been questioned why they did not become more important until the development of the inbred-line–hybrid concept in the 1920s. Three reasons are the following:

(1) The time was not right to realize the commercial potential for producing hybrid seed. Variety crosses were not significantly different in phenotypic appearance from varieties, and the heterosis was not as great as observed for hybrids produced among inbred lines.

(2) Research on inbreeding depression and heterosis among inbred lines was striking enough to divert attention from variety hybrids to the potential use of inbred lines to produce hybrids.

(3) Although comparisons showed that crosses were superior, the heterosis concept was not known, and the principal objective of these studies was to control the parentage of the seed being produced.

However, the extensive use of recurrent selection methods for population improvement and diallel mating designs later demonstrated that not only alternative heterotic
patterns could be proposed (Melani and Carena, 2005) but also population hybrids with similar performance to single-cross commercial hybrids could be identified (Carena, 2005).

The beginning of the heterosis concept in maize started with the studies reported by Shull (1908), ‘The composition of a field of maize.’ Shull (1952) summarized his studies and apparently was the first to correctly interpret the phenomena of inbreeding depression and hybrid vigor. Hybrid vigor and heterosis are nearly synonymous; the word heterosis was coined by Shull (1914) to provide a term to describe the phenomenon but it did not include a description of the genetic mechanism involved in its expression.

### 10.2 Empirical Evidence

The manifestation of heterosis in crosses of maize varieties by Beal (1880) alerted other researchers to possible benefits of variety crosses. Several studies evaluating open-pollinated varieties and their crosses in the early part of the 20th century were summarized by Richey (1922). Most of the studies involved crosses of a series of open-pollinated varieties with an open-pollinated variety that was either high yielding or popular with local farmers. Hence a group of open-pollinated varieties were crossed to a common tester variety to determine performance of the variety crosses and heterosis expressed in the variety crosses relative to either the average performance of the two open-pollinated parental varieties or the high parent. Because of Jones’s (1918) suggestion that single crosses be used as parent stocks for production of double-cross hybrids, procedures for inbred-line and hybrid development received strong emphasis during the 1920s; consequently, little information is available in the literature after 1920 for variety crosses. Performance of variety crosses received renewed interest in the 1950s because of the development of quantitative genetics and recurrent selection procedures for the improvement of breeding populations. Differences of opinion of relative importance of dominance vs. overdominance in expression of heterosis in hybrids, the apparent yield plateau in double-cross hybrid yields in the 1950s, and potential of recurrent selection procedures for improvement of breeding populations stimulated development of different maize breeding procedures. The apparent yield plateau suggested that adequate genetic variation was not present in open-pollinated varieties for development of superior double-cross hybrids. Quantitative genetic studies were dependent on a reference population for making inferences, and reciprocal recurrent selection was suggested on the premise of using two base populations and selecting for both general and specific combining abilities. In all instances the use of open-pollinated varieties became important because they were the source populations used to develop the original group of inbred lines for production and growing of double-cross hybrids as well as they were the source of improved populations.

The method of evaluation and the choice of varieties included for evaluation of heterosis also changed. Instead of crossing a group of varieties to a common tester
variety, the diallel mating design was used to determine general performance of a variety in comparison with other varieties and specific performance of a particular pair of varieties. The latter information was important in the choice of varieties and/or improved populations for initiating reciprocal recurrent selection (RRS). Open-pollinated varieties were included in many of the diallel series of crosses, but synthetic varieties, composites, and varieties improved by selection also were often included. In most instances a measure of heterosis was desired among the variety crosses, but in some instances genetic information was obtained by selfing either the parental varieties or the variety crosses.

Two methods were proposed to actually measure the performance of a hybrid relative to its parents:

(1) **Mid-parent (MP) heterosis (MPH)**: It is the performance of a hybrid relative to the average performance of its parents expressed in percentage.

\[ \text{MPH} = \frac{F_1 - \text{MP}}{\text{MP}} \times 100 \]

(2) **High-parent (HP) heterosis (HPH)**: It is the performance of a hybrid relative to the performance of its best parent expressed in percentage.

\[ \text{HPH} = \frac{F_1 - \text{HP}}{\text{HP}} \times 100 \]

The HP heterosis method has been less used but it provides better and more accurate information.

A summary of the manifestation of heterosis in crosses of maize varieties is given in Tables 10.1 and 10.2. Information on heterosis of crosses ranges from that of Morrow and Gardner (1893) to information evaluating effectiveness of recurrent selection. Because yield is the most important economic trait of maize, only the heterosis information on yield is given. Table 10.1 includes 611 varieties and 1394 variety crosses that were evaluated for yield heterosis. Note that many varieties and also some crosses are repeated in the numbers given in Table 10.1. Heterosis relative to the average of the two parent varieties (mid-parent) and the high-parent variety is given for each reported study and averaged over all studies. Average mid-parent heterosis for the 1394 crosses weighted for the number of crosses in each study was 19.5%. Average mid-parent heterosis was evident in nearly all studies; the only exception was for some of the varieties and variety crosses reported by Noll (1916), which was −0.5%. Mid-parent heterosis in Table 10.1 is the average for each study. Variety crosses that were either above or below the mid-parent also are shown in Table 10.1. Except for the study by Noll (1916) a majority of variety crosses exceeded the mid-parent values. High-parent heterosis and frequency of variety crosses that exceeded the high parent varied considerably among the reported studies. High-parent heterosis for variety crosses evaluated before 1932
<table>
<thead>
<tr>
<th>Source</th>
<th>Parent varieties</th>
<th>Variety crosses</th>
<th>Average heterosis (%)</th>
<th>Mid-parent (no.)</th>
<th>Exceeds high parent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mid-parent</td>
<td>High parent</td>
<td>Above</td>
</tr>
<tr>
<td>Morrow and Gardner (1893)</td>
<td>7</td>
<td>5</td>
<td>14.9</td>
<td>7.9</td>
<td>5</td>
</tr>
<tr>
<td>Hayes and East (1911)</td>
<td>3</td>
<td>2</td>
<td>41.2</td>
<td>5.0</td>
<td>2</td>
</tr>
<tr>
<td>Hartley et al. (1912)</td>
<td>81</td>
<td>75</td>
<td>9.8</td>
<td>−1.9</td>
<td>62</td>
</tr>
<tr>
<td>Hayes (1914)</td>
<td>39</td>
<td>35</td>
<td>10.7</td>
<td>−0.1</td>
<td>28</td>
</tr>
<tr>
<td>Williams and Welton (1915)</td>
<td>13</td>
<td>17</td>
<td>7.4</td>
<td>1.7</td>
<td>13</td>
</tr>
<tr>
<td>Noll (1916)</td>
<td>6</td>
<td>10</td>
<td>−0.5</td>
<td>−6.8</td>
<td>3</td>
</tr>
<tr>
<td>Hutchison and Wolfe (1917)</td>
<td>5</td>
<td>4</td>
<td>4.4</td>
<td>−9.0</td>
<td>3</td>
</tr>
<tr>
<td>Jones et al. (1917)</td>
<td>55</td>
<td>50</td>
<td>7.9</td>
<td>1.1</td>
<td>43</td>
</tr>
<tr>
<td>Hayes and Olson (1919)</td>
<td>13</td>
<td>12</td>
<td>14.3</td>
<td>11.5</td>
<td>12</td>
</tr>
<tr>
<td>Grifee (1922)</td>
<td>5</td>
<td>6</td>
<td>14.1</td>
<td>6.9</td>
<td>6</td>
</tr>
<tr>
<td>Kiesselbach (1922)</td>
<td>14</td>
<td>13</td>
<td>−3.6</td>
<td>−8.8</td>
<td>5</td>
</tr>
<tr>
<td>Waldron (1924)</td>
<td>13</td>
<td>12</td>
<td>9.1</td>
<td>4.8</td>
<td>11</td>
</tr>
<tr>
<td>Garber et al. (1926)</td>
<td>5</td>
<td>8</td>
<td>37.2</td>
<td>−2.1</td>
<td>8</td>
</tr>
<tr>
<td>Garber and North (1931) − 0</td>
<td>2</td>
<td>1</td>
<td>45.4</td>
<td>13.6</td>
<td>1</td>
</tr>
<tr>
<td>Garber and North (1931) − 1</td>
<td>2</td>
<td>1</td>
<td>32.9</td>
<td>−9.9</td>
<td>1</td>
</tr>
<tr>
<td>Robinson et al. (1956)</td>
<td>6</td>
<td>15</td>
<td>19.9</td>
<td>11.5</td>
<td>15</td>
</tr>
<tr>
<td>Torregroza (1959)</td>
<td>10</td>
<td>7</td>
<td>17.0</td>
<td>—</td>
<td>7</td>
</tr>
<tr>
<td>Lonnquist and Gardner (1961)</td>
<td>12</td>
<td>66</td>
<td>8.5</td>
<td>2.8</td>
<td>61</td>
</tr>
<tr>
<td>Moll et al. (1962)</td>
<td>6</td>
<td>14</td>
<td>20.2</td>
<td>8.6</td>
<td>13</td>
</tr>
<tr>
<td>Paterniani and Lonnquist (1963)</td>
<td>12</td>
<td>63</td>
<td>33.0</td>
<td>14.0</td>
<td>61</td>
</tr>
<tr>
<td>Timothy (1963)</td>
<td>8</td>
<td>28</td>
<td>31.0</td>
<td>22.0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>6</td>
<td>50.3</td>
<td>15.1</td>
<td>6</td>
</tr>
<tr>
<td>Moll et al. (1965)</td>
<td>8</td>
<td>28</td>
<td>22.8</td>
<td>8.7</td>
<td>28</td>
</tr>
<tr>
<td>Wellhausen (1965)</td>
<td>15</td>
<td>18</td>
<td>59.2</td>
<td>35.9</td>
<td>18</td>
</tr>
<tr>
<td>Hallauer and Eberhart (1966)</td>
<td>9</td>
<td>36</td>
<td>11.0</td>
<td>6.0</td>
<td>36</td>
</tr>
</tbody>
</table>
Table 10.1  (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Parent varieties</th>
<th>Variety crosses</th>
<th>Average heterosis (%)</th>
<th>Mid-parent (no.)</th>
<th>Exceeds high parent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mid-parent</td>
<td>High parent</td>
<td>Above</td>
</tr>
<tr>
<td>Paterniani (1967)</td>
<td>9</td>
<td>26</td>
<td>40.5</td>
<td>30.1</td>
<td>26</td>
</tr>
<tr>
<td>Castro et al. (1968)</td>
<td>5</td>
<td>10</td>
<td>45.5</td>
<td>2.9</td>
<td>10</td>
</tr>
<tr>
<td>Paterniani (1968)</td>
<td>10</td>
<td>45</td>
<td>34.8</td>
<td>5.6</td>
<td>43</td>
</tr>
<tr>
<td>Crum (1968)^b</td>
<td>10</td>
<td>45</td>
<td>21.3</td>
<td>7.0</td>
<td>43</td>
</tr>
<tr>
<td>Hallauer and Sears (1968)</td>
<td>9</td>
<td>36</td>
<td>9.8</td>
<td>4.2</td>
<td>35</td>
</tr>
<tr>
<td>Troyer and Hallauer (1968)</td>
<td>10</td>
<td>45</td>
<td>72.0</td>
<td>43.0</td>
<td>45</td>
</tr>
<tr>
<td>Silva (1969)</td>
<td>3</td>
<td>3</td>
<td>18.2</td>
<td>0.3</td>
<td>3</td>
</tr>
<tr>
<td>Paterniani (1970)</td>
<td>16</td>
<td>55</td>
<td>2.6</td>
<td>−5.6</td>
<td>33</td>
</tr>
<tr>
<td>Moll and Stuber (1971)^c−0</td>
<td>2</td>
<td>1</td>
<td>18.9</td>
<td>17.8</td>
<td>1</td>
</tr>
<tr>
<td>− I</td>
<td>2</td>
<td>2</td>
<td>23.0</td>
<td>21.0</td>
<td>2</td>
</tr>
<tr>
<td>Eberhart (1971) – Corn Belt</td>
<td>9</td>
<td>36</td>
<td>14.2</td>
<td>4.5</td>
<td>32</td>
</tr>
<tr>
<td>− Southern</td>
<td>6</td>
<td>15</td>
<td>21.4</td>
<td>11.2</td>
<td>15</td>
</tr>
<tr>
<td>Tavares (1972)</td>
<td>6</td>
<td>3</td>
<td>13.6</td>
<td>0.4</td>
<td>3</td>
</tr>
<tr>
<td>Hallauer (1972)</td>
<td>9</td>
<td>36</td>
<td>14.0</td>
<td>8.1</td>
<td>32</td>
</tr>
<tr>
<td>El-Rouby and Galal (1972)</td>
<td>7</td>
<td>21</td>
<td>6.4</td>
<td>−0.8</td>
<td>20</td>
</tr>
<tr>
<td>Crum (1972)^b</td>
<td>10</td>
<td>45</td>
<td>18.4</td>
<td>11.2</td>
<td>43</td>
</tr>
<tr>
<td>Barriga and Vencovsky (1973)</td>
<td>5</td>
<td>10</td>
<td>13.7</td>
<td>2.5</td>
<td>10</td>
</tr>
<tr>
<td>Valois (1973)^c−0</td>
<td>2</td>
<td>1</td>
<td>9.8</td>
<td>9.5</td>
<td>1</td>
</tr>
<tr>
<td>− I</td>
<td>2</td>
<td>1</td>
<td>11.6</td>
<td>8.1</td>
<td>1</td>
</tr>
<tr>
<td>− III</td>
<td>2</td>
<td>1</td>
<td>4.2</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>Vencovsky et al. (1973)</td>
<td>10</td>
<td>45</td>
<td>21.6</td>
<td>12.0</td>
<td>45</td>
</tr>
<tr>
<td>Miranda (1974a)</td>
<td>9</td>
<td>36</td>
<td>15.2</td>
<td>5.8</td>
<td>32</td>
</tr>
<tr>
<td>Genter and Eberhart (1974)^−0</td>
<td>6</td>
<td>15</td>
<td>15.7</td>
<td>10.8</td>
<td>15</td>
</tr>
<tr>
<td>− 0×I</td>
<td>13</td>
<td>42</td>
<td>14.9</td>
<td>9.0</td>
<td>42</td>
</tr>
<tr>
<td>− I</td>
<td>7</td>
<td>21</td>
<td>14.2</td>
<td>8.8</td>
<td>21</td>
</tr>
</tbody>
</table>
Table 10.1 (continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>Parent varieties</th>
<th>Variety crosses</th>
<th>Average heterosis (%)</th>
<th>Mid-parent (no.)</th>
<th>Exceeds high parent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Above</td>
<td>Below</td>
</tr>
<tr>
<td>Shehata – set 1</td>
<td>10</td>
<td>45</td>
<td>24.0</td>
<td>11.8</td>
<td>26</td>
</tr>
<tr>
<td>and Dhawan (1975) – set 2</td>
<td>10</td>
<td>45</td>
<td>19.0</td>
<td>8.0</td>
<td>22</td>
</tr>
<tr>
<td>– set 3</td>
<td>10</td>
<td>45</td>
<td>19.0</td>
<td>9.2</td>
<td>31</td>
</tr>
<tr>
<td>Hallauer – 0</td>
<td>7</td>
<td>21</td>
<td>18.8</td>
<td>5.8</td>
<td>21</td>
</tr>
<tr>
<td>And Malithano (1976)^c – I</td>
<td>10</td>
<td>45</td>
<td>20.6</td>
<td>12.2</td>
<td>45</td>
</tr>
<tr>
<td>Paterniani (1977)</td>
<td>12</td>
<td>36</td>
<td>9.7</td>
<td>2.8</td>
<td>30</td>
</tr>
<tr>
<td>Obilana et al. (1979) – 0</td>
<td>2</td>
<td>1</td>
<td>10.8</td>
<td>8.1</td>
<td>1</td>
</tr>
<tr>
<td>– I</td>
<td>2</td>
<td>1</td>
<td>16.5</td>
<td>13.0</td>
<td>1</td>
</tr>
<tr>
<td>Paterniani and Goodman (1977)</td>
<td>6</td>
<td>15</td>
<td>18.7</td>
<td>7.9</td>
<td>15</td>
</tr>
<tr>
<td>Totals and means^d</td>
<td>611</td>
<td>1394</td>
<td>19.5</td>
<td>8.2</td>
<td>1206</td>
</tr>
</tbody>
</table>

^aHeterosis of variety crosses was 16% or greater
^bPersonal communication
^cData for original (0) and improved (I) variety crosses also are included in Chapter 7
^dWeighted by the number of crosses
Table 10.2  Summary of comparisons between parental varieties of maize and the first-generation crosses for different eras and original and improved varieties

<table>
<thead>
<tr>
<th>Source</th>
<th>Parent varieties</th>
<th>Variety crosses</th>
<th>Average heterosis (%)</th>
<th>Mid-parent (%)</th>
<th>Exceeds high parent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mid-parent</td>
<td>High parent</td>
</tr>
<tr>
<td>Richey (1922)</td>
<td>—</td>
<td>244</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Table 10.1</td>
<td>611</td>
<td>1394</td>
<td>19.5</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>Before 1932</td>
<td>263</td>
<td>251</td>
<td>9.9</td>
<td>0.0&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>After 1955</td>
<td>348</td>
<td>1143</td>
<td>21.6</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>Original</td>
<td>21</td>
<td>40</td>
<td>17.9</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>25</td>
<td>71</td>
<td>18.8</td>
<td>11.1</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Greater than 16%

<sup>b</sup>Exact value is −0.028%
was generally quite small. Average high-parent heterosis ranged from −9.9% for the one variety cross reported by Garber and North (1931) to 43.0% for 10 flint variety crosses reported by Troyer and Hallauer (1968). Average high-parent heterosis for the 1394 variety crosses was 8.2%.

Mid-parent (MP) and high-parent (HP) heterosis values were gathered for 71 improved populations in the 1980s. The average MP heterosis across improved population crosses was 19.5%, while the average HP heterosis across the same population crosses was 8.2%. One of the reasons variety crosses were not widely accepted is because choice of germplasm sources for inbred lines and their improve versions was not ideal. Weatherspoon (1973) suggested that in order for recurrent selection to be successful the initial germplasm pool should be the most elite material available. A more careful selection of improved germplasm after extensive testing can improve average values of mid- and high-parent heterosis to 38.9 and 28.2%, respectively.

Table 10.2 gives a summary of mid-parent and high-parent heterosis for six different categories of variety crosses. Compared with the summary of heterosis prepared by Richey (1922), percentages of variety crosses that exceeded the mid-parent and high parent were amazingly similar. Table 10.1 also includes some of those reported by Richey, but only 244 of the 1394 crosses were available at the time of Richey’s summary. The percentages of the variety crosses that exceeded the high parent by 0–5, 6–15, and 16% or more also were very similar. The most obvious difference of the comparisons in Table 10.2 is between average high-parent heterosis for variety crosses tested before 1932 (0.0%) and those tested after 1955 (10.0%). Perhaps the small and inconsistent expression of high-parent heterosis was an important factor in the limited use of variety crosses before development of double crosses from use of inbred lines. Because high-parent heterosis was not always observed (variety crosses exceeded high-parent in only 53.0% of the crosses, Table 10.2) and often less than 5% (24.3% of the crosses, Table 10.2), it is not surprising that variety crosses were not widely accepted. The superiority of variety crosses generally was not as great as reported by Beal (1880). Often (as mentioned previously) varieties were crossed to a common tester. A highly productive variety used as tester also would have reduced expression of high-parent heterosis. Table 10.2 shows that 17.2% of variety crosses tested before 1932 yielded less than the mid-parent. Since 1955, 10.0% of variety crosses yielded less than the mid-parent; the diallel mating design was used in nearly all instances for the latter comparisons. Comparisons of the varieties and variety crosses for the populations included in the recurrent selection programs show that all the variety crosses, both original and improved, exceeded the mid-parent. This may be expected because in most instances the varieties included were based on previous variety cross information and specifically selected for heterosis manifested in variety crosses. Mid-parent heterosis and high-parent heterosis, however, were very similar for the original (8.3%) and improved variety crosses (11.1%).

As expected, there was considerable variation among individual variety crosses within each of the studies reported in Table 10.1. In the earlier studies, Richey (1922) pointed out that greatest heterosis was manifested in variety crosses of
extreme types. For instance, heterosis in crosses between flint or flour varieties and dent varieties was greater than crosses among dent varieties. No critical data are available but it seems that greatest expression of heterosis is not dependent on differences in only endosperm types. Paterniani and Lonnquist (1963) included different endosperm-type varieties and found that manifested heterosis was as great among variety crosses produced from varieties having the same endosperm type as among those produced from varieties having different endosperm types. Heterosis manifested in variety crosses was dependent on more than just endosperm type. Greatest expression of heterosis was reported by Troyer and Hallauer (1968) for crosses produced among early flint varieties.

Table 7.19 (Chapter 7) compares the mid-parent and high-parent heterosis values between early studies on selected variety crosses (Tables 10.1 and 10.2) and later studies resulting from long-term recurrent selection programs. The number of successful heterotic combinations currently available is limited because public genetic diversity in reserve has not been fully exploited (Carena and Wicks III, 2006; Smith, 2007). Definitively, improvement of populations has increased heterosis values.

Two statements in summaries involving variety crosses illustrate the status of the potential of variety crosses. Hayes and Olson (1919) stated, ‘The use of first-generation crosses between pure varieties is a means of increasing yield of corn although all such crosses are not equally productive, some being of no value.’ Richey (1922) in his summary of variety crosses stated, ‘In such more or less haphazard crossing, therefore, the chances seem about equal of obtaining a cross that is or is not better than the better parent.’ The conclusions of these authors show that advantages of variety crosses were not always obvious. All the data were collected before experimental plot techniques and statistical analysis were developed, both of which would tend to mask the true differences. These conclusions were also made at the time vigorous attention was being given to the inbred-line–hybrid breeding programs. Heterosis might have been a tool used to focus research on the public ‘inbred–hybrid concept’ as the only mechanism for maize genetic improvement (Carena and Wicks III, 2006). As a consequence, efforts to improve varieties and their crosses were minimal. Renewed interest in improving breeding populations redirected some of the breeding efforts toward improving elite and genetically broad-based populations via recurrent selection. These efforts included extensive improvement of populations per se but limited testing of elite population crosses between divergent populations that have undergone long-term selection.

The manifestation of heterosis usually depends on genetic divergence of the two parental varieties. Genetic divergence among varieties usually is unknown, and the only recourse is to determine level of genetic divergence empirically by means of variety crosses. Genetic divergence of the parental varieties is inferred from the heterotic patterns manifested in the series of variety crosses. If heterosis manifested from the cross of two parental varieties is relatively large, it is concluded that the two parental varieties are more genetically diverse than two varieties that manifest little or no heterosis in their variety crosses. Establishment of heterotic patterns
among varieties has important implications for selecting inbred lines as potential seed stocks in hybrids (Hallauer et al., 1988). One of the first decisions maize breeders usually make in determining the matrix of crosses to produce among a set of elite inbred lines is the parental origin of the lines. If the origins of the inbred lines are known, the logical crosses can be produced based on the heterotic pattern of the parental source populations. For example, in the US Corn Belt, elite inbred lines of Reid Yellow Dent or BSSS background tend to be crossed and tested in combination with elite lines of Lancaster background. Although we have discussed importance of genetic divergence of elite inbred lines (Chapter 8), the same procedures are used for early testing of inbred lines to obtain preliminary information on their combining abilities. If origins of the lines are unknown, crosses usually are made to elite inbred lines of known origin. Subsequent yield test information provides the information necessary for classification into known heterotic patterns.

Examples for determining heterotic pattern of varieties and importance of genetic diversity in the manifestation of heterosis were reported by Moll et al. (1962), Moll et al. (1965), Tsotsis (1972), Kauffmann et al. (1982), Carena (2005), and Melani and Carena (2005). The objective of Tsotsis’s study was to determine the heterotic pattern among a group of open-pollinated varieties in order to establish two broad genetic base breeding populations. Because applied maize breeders want to maximize heterotic response in crosses among elite inbred lines, the task will be enhanced if genetic differences are established in the basic breeding populations. From evaluation of variety crosses Tsotsis was able to assign varieties to the two gene pools, based on the heterotic patterns of the variety crosses. The logical sequence is then to cross and test lines between the two gene pools. Tables 10.3 and 10.4 show a couple of examples targeted at identifying potential heterotic combinations. Note that low parental means can be misleading when measuring mid-parent heterosis values.

From Tables 10.3 and 10.4 alternative heterotic patterns were proposed (Kauffman et al., 1982; Melani and Carena, 2005), and intra- and inter-population recurrent selection programs were initiated (Carena and Hallauer, 2001; Hallauer and Carena, 2009).

Table 10.3  Yields (Mg ha\(^{-1}\)) of five US open-pollinated (O.P.) maize varieties (diagonal) and their diallel crosses (above diagonal) at two locations in 1970

<table>
<thead>
<tr>
<th>Varieties</th>
<th>Midland</th>
<th>Leaming</th>
<th>Lancaster</th>
<th>Reid</th>
<th>Krug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midland</td>
<td>6.04</td>
<td>6.43</td>
<td>6.99</td>
<td>6.22</td>
<td>6.21</td>
</tr>
<tr>
<td>Leaming</td>
<td>121.60(a)</td>
<td>4.54</td>
<td>6.05</td>
<td>4.85</td>
<td>5.88</td>
</tr>
<tr>
<td>Lancaster</td>
<td>132.60</td>
<td>133.85</td>
<td>4.49</td>
<td>5.79</td>
<td>5.71</td>
</tr>
<tr>
<td>Reid</td>
<td>107.43</td>
<td>96.23</td>
<td>115.57</td>
<td>5.53</td>
<td>4.63</td>
</tr>
<tr>
<td>Krug</td>
<td>112.30</td>
<td>122.50</td>
<td>118.96</td>
<td>87.36</td>
<td>5.02</td>
</tr>
</tbody>
</table>

\(a\)Mid-parent heterosis estimates are below diagonal (%)

Source: Adapted from Kauffman et al. (1982) and Hallauer et al. (1988)
Table 10.4 Yields (Mg ha\(^{-1}\)) of 10 short-season maize populations and their diallel crosses (above diagonal) at 29 environments

<table>
<thead>
<tr>
<th>Variety Cross</th>
<th>BS-5</th>
<th>BS-5/1(R)</th>
<th>BS-5/21(R)</th>
<th>LEAMING(S)</th>
<th>BS-22(R)</th>
<th>NDSM(M)</th>
<th>NDSG(M)</th>
<th>CGSS(S1-S2)</th>
<th>CGL(S1-S2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDSAB(MER)C12</td>
<td>5.60</td>
<td>5.87</td>
<td>7.27</td>
<td>5.58</td>
<td>5.88</td>
<td>6.36</td>
<td>6.08</td>
<td>5.73</td>
<td>5.99</td>
</tr>
<tr>
<td>BS-5</td>
<td>121</td>
<td>4.11</td>
<td>5.48</td>
<td>5.42</td>
<td>4.85</td>
<td>4.43</td>
<td>4.54</td>
<td>4.45</td>
<td>4.76</td>
</tr>
<tr>
<td>BS-21(R)C7</td>
<td>137</td>
<td>120</td>
<td>5.04</td>
<td>6.98</td>
<td>6.54</td>
<td>7.43</td>
<td>6.73</td>
<td>6.55</td>
<td>7.95</td>
</tr>
<tr>
<td>NDSG(M)C15</td>
<td>111</td>
<td>117</td>
<td>129</td>
<td>110</td>
<td>128</td>
<td>5.39</td>
<td>4.75</td>
<td>4.57</td>
<td>5.68</td>
</tr>
<tr>
<td>NDSG(M)C15</td>
<td>119</td>
<td>109</td>
<td>144</td>
<td>103</td>
<td>107</td>
<td>129</td>
<td>101</td>
<td>4.06</td>
<td>5.23</td>
</tr>
<tr>
<td>CGSS(S1-S2)C5</td>
<td>107</td>
<td>119</td>
<td>150</td>
<td>109</td>
<td>127</td>
<td>139</td>
<td>104</td>
<td>109</td>
<td>5.55</td>
</tr>
<tr>
<td>CGL(S1-S2)C5</td>
<td>116</td>
<td>113</td>
<td>156</td>
<td>108</td>
<td>141</td>
<td>139</td>
<td>100</td>
<td>124</td>
<td>119</td>
</tr>
</tbody>
</table>

Estimates of MP heterosis are below diagonal (mid-parent value + 100)
Hybrids in blue are not statistically different from the top commercial hybrid
Source: Adapted from Carena (2005)

Diallel cross analysis for a fixed set of open-pollinated varieties provides the basis for a preliminary analysis of the heterotic pattern among variety crosses (Gardner and Eberhart, 1966). Preliminary inferences are taken from the significance of effects in the analysis of variance. Thus, average heterosis \( \bar{h} \) is indicative of the superiority of variety crosses over mid-parent values. Variety heterosis, when significant as a source of variation, indicates that the heterotic pattern of at least one of the varieties differs from the others when crossed with the remaining varieties. Specific heterosis results from specific crosses, and when significant, it means that at least one cross differs from the others due to non-additive effects and differences in gene frequency of varieties. A summary of several diallel cross analyses for variety crosses relative to the significance of effects is shown in Table 10.5.

Only 4 out of 15 or 26.7% of the studies detected significance for specific effects for yield, suggesting that in most instances the choice of parental varieties for heterosis exploitation can be based on the yield of varieties themselves and on the average of their crosses. Only occasionally is the superiority of a given cross increased by positive and significant specific effects.

Moll et al. (1962, 1965) studied crosses produced among varieties with different levels of genetic diversity. The level of genetic diversity was inferred from the geographical origin of the varieties. Two open-pollinated varieties from the US Corn Belt, for instance, were assumed to have less genetic divergence than a variety from the US Corn Belt compared with one from southeastern United States. Moll et al. (1962) crossed in diallel series six open-pollinated varieties: two from the US Corn Belt, two from southeastern United States, and two that originated in the Caribbean region. All parental varieties and 14 of the possible 15 variety crosses were tested in North Carolina. Relative heterosis of the variety crosses agreed with the original classification of the genetic divergence of the parental varieties. Moll et al. (1965)
Table 10.5  Level of significance of variety and heterotic component effects for yield in several variety diallel crosses

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of varieties</th>
<th>Effects&lt;sup&gt;a&lt;/sup&gt;</th>
<th>v&lt;sub&gt;j&lt;/sub&gt;</th>
<th>h</th>
<th>h&lt;sub&gt;j&lt;/sub&gt;</th>
<th>s&lt;sub&gt;jj&lt;/sub&gt;′</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gardner (1965)</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Gardner and Eberhart (1966)</td>
<td>6</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Hallauer and Eberhart (1966)</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Gardner and Paterniani (1967)</td>
<td>6</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Hallauer and Sears (1968)</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Troyer and Hallauer (1968)</td>
<td>10</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Castro et al. (1968)</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Vencovsky (1969)</td>
<td>12</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Eberhart (1971) – Corn Belt</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>– Southern</td>
<td>6</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Hallauer (1972)</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Barriga and Vencovsky (1973)</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Miranda (1974a) – set 1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>– set 2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Genter and Eberhart (1974)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>13</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>ns</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Significant, P < 0.05; ns, non-significant
<sup>b</sup>See Section 10.6 for the meaning of effects
<sup>c</sup>Set 1: seven short plant varieties; set 2: set 1 plus two high-yielding varieties
<sup>d</sup>Includes six original (0) and seven advanced (I) populations; significance levels refer to each effect on the average of 0 and I; only h<sub>j</sub> showed a significant interaction with (0 vs. I)

extended the study to include two races from Mexico. The eight parental varieties and 28 variety crosses were tested in each region of origin of the parental varieties. They found that heterosis manifested in crosses of varieties hypothesized to be the most genetically diverse was less than heterosis expressed between varieties considered to be less genetically diverse. These results suggested that the concept of genetic divergence for maximum expression of heterosis has limits. Apparently crosses between extremely divergent parents create a situation where the harmonious functioning of alleles is disrupted. Consequently the physiological functions are not as efficient as in situations where the alleles have had similar selection pressures. Several years later we have also learned recurrent selection can affect the genetic divergence produced by geographical origin.

Heterosis observed in variety crosses is the average expression of heterosis of the genotypes formed by crossing a sample of genotypes from each of the two parental varieties. Individual F<sub>1</sub> plants in the variety crosses will vary in relative heterosis, depending on the two parental gametes sampled to produce the F<sub>1</sub> plant. Some of the F<sub>1</sub> variety cross plants would be expected to show considerably more heterosis than others. If sampling is adequate the range of heterosis of individual plants in the variety crosses will approximate a normal distribution. The primary goal is to identify the parental plants in each variety that maximize heterosis when crossed (e.g., the performance of a specific hybrid is greater than the modal performance of
a variety cross). This concept was the basis of Shull’s (1908) suggestion for developing pure lines and using them to produce the crosses. Not all crosses among a set of lines will exceed the variety cross. If a set of unselected inbred lines is developed from a variety and a large number of crosses is made among the unselected inbred lines, the average of all crosses is expected to regenerate the original variety (Good and Hallauer, 1977). Since Shull’s suggestion, the objective of maize breeding has been to obtain an elite group of inbred lines and to identify the specific set of lines that maximizes expression of heterosis in hybrids. Because the inbreds are nearly homogeneous and homozygous, the superior hybrid identified is reproducible. Identified population hybrids from improved populations could also be reproducible if large samples are maintained. Consequently, expression of heterosis for yield in say single-cross hybrids of inbred lines would be expected to be much greater.

One example is provided by data of Martin and Hallauer (1976) (Table 10.6). They compared means of four sets of inbred lines and their single-cross hybrids in the same experiment. Heterosis for yield exceeded 150% for each of the four sets of inbred lines. Heterosis manifested in single crosses produced from selected homozygous lines is about ten-fold greater than in varietal crosses shown in Tables 10.1 and 10.2.

Differences in magnitudes of heterosis observed between the two sets of material have at least two obvious explanations:

(1) The base of comparison is much different in most instances because average yield of two inbred lines is less than that of two non-inbred varieties.
(2) Inbred lines included by Martin and Hallauer were selected because of their performance in hybrids and would not be expected to represent a random sample of gametes from a maize population.

The first explanation is common in comparisons that involve means. The greater heterosis among flint variety crosses reported by Troyer and Hallauer (1968) in comparison with dent varieties studied by Lonnquist and Gardner (1961) could be explained in part by differences in yield levels of parents included and environments in which they were tested. The second explanation is a consequence of the breeding procedures used to develop elite inbred lines for use in single-cross hybrids.

We have learned that heterosis is dependent on genetic divergence in the presence of dominance (Moll et al., 1965; Falconer and Mackay, 1996). As a consequence, the heterotic response was especially improved by continuous selection for specific combining ability (Eyherabide and Hallauer, 1991; Keeratinijakal and Lamkey, 1993; Labate et al., 1997; Menz et al., 1999). Accordingly, reciprocal recurrent selection programs with, on an average, over eight selection cycles were effective for improving the grain yield performance of population hybrids. In these programs, high-parent heterosis values ranged from 22.5 to 72.4 %. Moreover, improved populations were fivefold higher in heterosis than their respective unimproved versions, showing the importance of selection in heterosis expression. In addition, Carena
### Table 10.6 Heterosis for four groups of inbred lines for yield and four components of yield

<table>
<thead>
<tr>
<th>Generation</th>
<th>Trait</th>
<th>Yield (g/plant)</th>
<th>Ear length (cm)</th>
<th>Ear diameter (cm)</th>
<th>Kernel-row number</th>
<th>300-kernel wt. (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mid-parent</td>
<td>F$<em>{2}$, BC$</em>{1}$, BC$_{2}$</td>
<td>F$_{1}$</td>
<td>$H^a$</td>
<td>$I^a$</td>
</tr>
<tr>
<td>Group I</td>
<td></td>
<td>48.62</td>
<td>12.60</td>
<td>3.82</td>
<td>14.16</td>
<td>63.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100.97</td>
<td>15.87</td>
<td>4.30</td>
<td>15.24</td>
<td>70.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>152.65</td>
<td>18.37</td>
<td>4.57</td>
<td>15.84</td>
<td>76.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>213.96</td>
<td>45.8</td>
<td>19.6</td>
<td>11.9</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−33.85</td>
<td>−13.6</td>
<td>−5.9</td>
<td>−3.8</td>
<td>−7.2</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td>63.58</td>
<td>14.68</td>
<td>4.02</td>
<td>14.46</td>
<td>71.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>116.36</td>
<td>16.64</td>
<td>4.46</td>
<td>15.45</td>
<td>75.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>159.59</td>
<td>18.19</td>
<td>4.71</td>
<td>15.78</td>
<td>82.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>151.00</td>
<td>23.9</td>
<td>17.2</td>
<td>9.1</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−27.08</td>
<td>−8.5</td>
<td>−5.3</td>
<td>−2.1</td>
<td>−8.2</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td>67.04</td>
<td>14.67</td>
<td>4.04</td>
<td>14.73</td>
<td>67.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>111.21</td>
<td>16.45</td>
<td>4.40</td>
<td>15.80</td>
<td>68.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150.72</td>
<td>18.22</td>
<td>4.62</td>
<td>16.41</td>
<td>72.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>124.82</td>
<td>24.2</td>
<td>14.4</td>
<td>11.4</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−26.22</td>
<td>−9.7</td>
<td>−4.8</td>
<td>−3.7</td>
<td>−5.1</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td>38.24</td>
<td>12.73</td>
<td>3.67</td>
<td>13.26</td>
<td>73.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>103.59</td>
<td>16.58</td>
<td>4.27</td>
<td>15.12</td>
<td>72.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>158.47</td>
<td>19.25</td>
<td>4.55</td>
<td>15.63</td>
<td>76.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>314.4</td>
<td>51.2</td>
<td>24.0</td>
<td>17.9</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−34.64</td>
<td>−13.9</td>
<td>−6.2</td>
<td>−3.3</td>
<td>−4.6</td>
</tr>
</tbody>
</table>

$aH$ (heterosis) = \[\frac{(F_1-\text{MP})}{\text{MP}} \times 100; \]

$I$ (inbreeding depression) = \[\frac{(F_1-F_2)}{F_1} \times 100\]

Source: Martin and Hallauer (1976)

(2005) presented evidence that crosses between geographically isolated populations improved by both intra- and inter-population selection programs also can identify outstanding population hybrids with significantly greater heterosis expression. Additive genetic effects accounted for the most successful crosses tested in North Dakota (Melani and Carena, 2005) without the need for inter-population selection programs (e.g., extensive testing can identify these crosses). Lonnquist (1963) suggested that breeding procedures have not made full use of the additive and non-additive genetic variation existent in elite populations.

Maize breeders should increase the number of successful heterotic combinations (Duvick, 1981). Most of the US maize hybrid industry is based upon the utilization of few heterotic patterns (Iowa Stiff Stalk Synthetic × Lancaster Sure Crop, Iowa Stiff Stalk Synthetic × Iodent) and the identification of alternative heterotic patterns should be emphasized. If greater emphasis is given to diverse improved populations, alternative and productive heterotic patterns will be identified.
We have only discussed observed heterosis for yield. Several studies have been reported attempting to define the origin of heterosis and its causes. These studies included different plant organs at different stages of development of parents and crosses. Generally they did not give an adequate explanation of the phenomenon of heterosis. Measurements usually showed that differences existed, but the basic mechanism was not determined. For a detailed account of these studies for maize and other plant species, the reader is referred to the summary given by Sprague (1953).

Significant investment on public laboratories across the nation has generated genotypic information with molecular markers as well as the recent genome sequence of B73. As heterotic effects are specific to each hybrid it has been extremely challenging to understand the molecular basis for heterosis and propose more convincing explanations than the ones already in place against all predictions published in the past 10 years.

10.3 Genetic Basis

Heterosis, whether between crosses of varieties or inbred lines, is observed in maize, but the genetic basis of observed heterosis is still conjectural (Coors and Pandey, 1999). Different theories have been proposed as an explanation, some having and others not having a genetic basis. Although extensive research has been conducted for the past 100 years, it has been difficult to prove or disprove the different theories proposed. Because of the importance of maize and the heterosis manifested in maize crosses, breeders and geneticists have been active in studying the heterosis phenomenon and attempting to develop genetic models for its basis.

Several hypotheses have tried to explain the causes of heterosis based upon the relationship between the level of dominance and the expression of heterosis without success. The existing data support the dominance hypothesis as the genetic basis of heterosis (being caused by the accumulation of favorable alleles with partial to complete dominance). Most of the hypotheses that have been proposed and discussed to account for heterosis in maize can be included in either one of the two following categories:

(a) Physiologic stimulation (or allelic interaction or overdominance).
(b) Dominant favorable growth factors.

The evidence supporting either of the two hypotheses depends on the available data and the interpretation of these data. Since the hypotheses were proposed and discussed, rapid progress has been made in quantitative genetic theory. Yield, as an example, has a very complex mode of inheritance, and heterosis of yield has been of great interest to both early and recent researchers. Yield is a measure of reproductive capacity and nearly always is treated as a quantitative trait. Information from quantitative genetic studies has, therefore, contributed to the understanding of heterosis.
in maize. In the past, however (e.g., see Richey, 1946), heterosis and quantitative genetics were often treated as separate entities, which they are not. Reviews on heterosis explanations were discussed by Whaley (1944), Richey (1946), Sprague (1953), and Coors and Pandey (1999).

Shull (1908) presented the first theory of heterosis, designated as the physiologic stimulation or heterozygosity hypothesis. The theory was based on the premise that heterozygosity itself was the cause of heterosis, which is a non-Mendelian explanation. East and Hayes (1912) and Shull (1912) supported this theory. In 1936, East reviewed the evidence for heterosis and concluded that dominant favorable growth factors were inadequate to explain heterosis. He proposed that multiple alleles at a locus are differentiated with respect to their physiological functions. As Sprague (1953) indicated, the ideas proposed are included within the limits of the physiologic stimulation hypothesis, but the proposed model was similar to the concept of overdominance suggested by Hull (1945).

The second hypothesis of heterosis, dominant favorable growth factors, was first presented mathematically by Bruce (1910). Although the hypothesis was briefly presented, the mathematical derivation given by Bruce is amenable to the current status of the dominance hypothesis and the occurrence and magnitude of heterosis that can be expected from crosses. The derivations presented by Bruce could include any number of gene pairs, any range of gene frequencies, and any level of dominance. The salient features of Bruce’s hypothesis were that heterosis would occur if the parents differed in gene frequency and dominance was present. It was a Mendelian genetic hypothesis. Jones (1917, 1945, 1958), among others, was a strong supporter of the hypothesis that heterosis is manifested by the accumulation of dominant favorable factors at different loci.

Proponents of the overdominance hypothesis attacked the dominance hypothesis because evidence did not substantiate the proposed model due to the following reasons:

1. If the phenomenon of heterosis were due to the accumulation of dominant favorable growth, one should be able to obtain inbred lines as productive as the single-cross hybrids and this has never been accomplished.

A response to this critic would be that although critical data are not available, empirical evidence indicates that relative vigor, health, and productivity of inbred lines are increasing in successive line recycling, which is why it is possible to produce single-cross hybrids in the US Corn Belt (Duvick, 1999). Some evidence is given in Table 10.6. Second-cycle inbred lines were 30.8% greater yielding than first-cycle lines. Also, Collins (1921) showed that if number of factor pairs for a trait exceeds 10, the chance of obtaining a plant homozygous for all 10 factors is remote. Gene number for traits such as vigor, health, and productivity is unknown but they must be a large number in all instances. Resolution of gene structure indicates that the nucleotide base pairs influencing yield are extremely large. Genes defined on this basis (e.g., QTL for yield) would preclude any opportunities of obtaining inbred lines as vigorous as single-cross hybrids.
But advances are being made in improvement of vigor, health, and productivity of our inbred lines and would substantiate the dominance hypothesis.

2. Absence of skewed distributions in the F2 populations also is an indication that dominance is not the primary feature of heterosis. If dominance were present, we would expect a skewed distribution from the expansion of the binomial \((3 + 1)^n\). Collins (1921) showed, however, that skewness because of dominance is not great if a large number of factor pairs are involved in expression of the trait. Again, for yield where number of factor pairs must be extremely large, absence of skewed distributions does not detract from the dominance hypothesis as the basis of heterosis.

3. Failure to obtain convincing evidence for occurrence of dominance in the expression of traits quantitatively inherited also is given as a reason for not supporting the dominance hypothesis.

On the contrary, most evidence indicates that partial to complete dominance is the primary mode of inheritance. Most evidence indicates overdominance is not the primary mode of inheritance. Some of the first quantitative genetic studies of F2 populations indicated that overdominance (see Chapter 5) may be important. After several rounds of recombination it was found, however, that estimates were biased because of linkage, a common feature of F2 populations. Thus, pseudo-overdominance because of repulsion phase linkages rather than true overdominance existed. Also, results of recurrent selection experiments designed to maximize selection for overdominant gene action (Hull, 1945) show that partial to complete dominance is more in concert with results of selection than overdominant gene action. Level of dominance is difficult to estimate for complex traits controlled by a large unknown number of genes, but most of the cumulative evidence favors partial to complete dominance for the majority of genes involved. One exception was a study by Horner et al. (1989), who concluded that overdominance effects were important for greater response to half-sib selection vs. inbred progeny selection.

Definitive proof for either of the hypotheses proposed for the genetic basis of heterosis probably will be difficult to establish. Because of complexity of the inheritance of quantitative traits, all types of gene action, both inter- and intra-allelic, are probably involved. Maize breeders and geneticists work on the basis of chromosome segments, and the resolution of factor pairs is slowly improving while the cost of genotyping keeps reducing. However, heterosis in maize can still be utilized without knowledge of the exact genetic basis of its occurrence. What is important is knowledge of the predominant types of gene action operative in designing effective and efficient breeding schemes for continued progress. For practical purposes, it seems that evidence supports the hypothesis that heterosis results from an accumulation of dominant favorable growth factors and finding ways to exploit them seems to be an answer to improve the expression of heterosis.

Crow (1948, 1952) presented arguments that the dominance hypothesis advanced by Bruce (1910) and Jones (1917) was not sufficient to account for observed heterosis among hybrids obtained by crossing inbred lines extracted from an equilibrium
population. Crow’s argument was based on mathematical calculations for a given set of assumptions, some of which are not valid in applied breeding programs. Crow defined an equilibrium population as one in which gene frequencies were in equilibrium between mutation and selection, and genotypic frequencies were those expected with random mating and linkage equilibrium. With this definition of an equilibrium population and assuming 5,000 loci and $10^{-5}$ mean mutation rate as reasonable estimates for yield it was calculated that maximum heterosis manifested in hybrids over the parent equilibrium population would be 5%. Hence, the best hybrid among inbred lines extracted from the equilibrium population would not exceed the equilibrium by more than 5%. Because the best hybrids among a set of inbred lines usually exceeds the equilibrium population by more than 5%, the only plausible explanation of heterosis was the importance of overdominance, as suggested by East (1936) and Hull (1945). The critical feature of Crow’s theoretical derivation is the definition of an equilibrium population. To maximize heterosis manifested in crosses of inbred lines, maize breeders usually cross lines between populations that show evidence of a definite heterotic pattern, such as lines of Reid Yellow Dent origin with those of Lancaster Surecrop origin or dent with flint. If the lines were extracted from different varieties, the assumption of an equilibrium population would apply only if all varieties are genetically identical, as in the case of gene frequencies being the same for all source populations from which the lines were extracted. Evidence given in Tables 10.1 and 10.2, however, demonstrates genetic differences among open-pollinated varieties because of the heterosis manifested among the variety crosses. It is interesting that the 5% derived by Crow is similar to the average high parent of 8.2% given in Table 10.1. If we visualize the maize species *Zea mays* L. as our equilibrium population, the estimate of 5% is similar to the observed high-parent heterosis of 8.2%. If the number of gene loci affecting yield is greater than 5,000 the average high-parent heterosis would satisfy Crow’s model. Unless the open-pollinated varieties have been selected for a common equilibrium status, differences in average grain yield between parental varieties and selected hybrids involving inbred lines extracted from different source varieties do not provide critical evidence relative to the level of dominance. But heterosis manifested in crosses among open-pollinated varieties (Tables 10.1 and 10.2) provides evidence that genetic differences exist among varieties. Usually, maize breeders produce hybrids between lines of diverse parentage. Because there are genetic differences among source populations or varieties, Crow’s arguments do not negate the dominance hypothesis as an explanation of heterosis in hybrids produced among inbred lines, or, for that matter, variety crosses produced among different varieties.

Epistasis also may contribute to the heterosis expressed in crosses. Although studies indicate epistasis does not seem to be a major component of genetic variability (see Chapter 5), epistatic effects have been shown to occur in specific crosses of inbred lines of maize (Bauman, 1959; Gorsline, 1961; Gamble, 1962; Sprague et al., 1962; Sprague and Thomas, 1967; Eberhart and Hallauer, 1968; Stuber and Moll, 1971; Moreno-Gonzalez and Dudley, 1981; Dudley and Johnson, 2009, et c). Detection of epistatic effects indicates that some specific crosses of lines with
unique combinations of genes contribute to heterosis. Cress (1966) has shown that multiple alleles can show negative dominance effects among some of the combinations, and hence could account for the observed results in the absence of epistasis. The curvilinear relation among different generations generated from crosses of inbred lines and varieties also is interpreted as evidence of epistasis (see Chapter 9). Epistasis for quantitative traits surely exists, but it has been difficult to determine that epistatic interactions account for very little of the genetic variability in maize populations beyond that accounted for by additive and dominance variances. Although it has not been quantified, epistasis may be important in the expression of heterosis in a single cross of two inbred lines as the unique combination of gene interactions is restricted to the cross of the two inbred lines. The possible role of epistatic effects, however, in expression of heterosis cannot be ignored (Coors and Pandey, 1999; Melchinger et al. 2007). Dudley and Johnson (2009) also reported that with the inclusion of epistasis in their prediction models that correlations between predicted and observed means were high enough to suggest they could be useful in maize breeding.

10.4 Biometrical Concept

As introduced before heterosis (H) is expressed when the parents of a hybrid have different alleles at a locus and there is some level of dominance among those alleles (Falconer and Mackay, 1996):

\[ H = \sum y^2 d \]

being \( d \) = dominance and \( y^2 \) = difference in allele frequencies between parents

If \( d = 0 \), the difference in allelic frequencies will not, in theory, contribute to heterosis.

Therefore, Falconer and Mackay (1996) has shown that heterosis will be expressed when we have the following conditions:

(1) Presence of some level of dominance.
(2) Relative difference in gene frequency of the two parents to determine the magnitude of the heterosis expressed in crosses.

If either or both of the conditions do not exist, heterosis will not be manifested. The conditions demonstrated by Falconer are essentially the same as those given by Bruce (1910), but they are generalized to include any number of genes, range in gene frequency of the two parents, and arbitrary level of dominance. The statistical description of heterosis given by Falconer does not support or detract from either of the proposed hypotheses for heterosis. As previously shown, heterosis is defined as \( H = \bar{F}_1 - \bar{M}P \), where \( \bar{F}_1 \) is the mean of the first-generation cross between two
parental populations of lines in Hardy–Weinberg equilibrium and \(\overline{MP}\) is the mid-parent value.

Chapter 2 shows that at the one locus level the population mean is

\[
\mu_A = (p - q)a + 2pqd
\]

where \(p\) and \(q = 1 - p\) refer to gene frequencies and \(a\) and \(d\) to genotype effects.

If another population is considered, with corresponding gene frequencies \(r\) and \(s\), its mean is given by

\[
\mu_B = (r - s)a + 2rsd
\]

Thus, the mid-parent value is

\[
\overline{MP} = \left(\frac{1}{2}\right)(p - q + r - s)a + (pq + rs)d
\]

From Chapter 2 we obtained the mean of the first-generation cross between two populations:

\[
F_1 = (pr - qs)a + (ps + qr)d
\]

Therefore, we have:

\[
H = [(pr - qs) - \left(\frac{1}{2}\right)(p - q + r - s)a + [(ps + qr) - (pq - rs)]d
\]

\[
= \left(\frac{1}{2}\right)[2pr - 2qs - (p - q) + (r - s)a + (ps + qr - pq - rs)d
\]

\[
= \left(\frac{1}{2}\right)(q^2 - s^2 - 2qs - p^2 + r^2 + 2rp)a + p(s - q)d + r(q - s)d
\]

\[
= \left(\frac{1}{2}\right)[(q - s)^2 - (p - r)^2]a + (s - q)(p - r)d
\]

\[
= 0 + (1 - r - 1 + p)(p - r)d
\]

\[
= (p - r)^2d
\]

which was expressed by Falconer and Mackay (1996) as

\[
H = \sum y^2d
\]

being the summation over all loci with \(y\) being difference in allele frequencies.

From the above derivations, it can be seen that the crossed population mean equals the mid-parent value when no dominance is assumed.

The \(F_2\) generation mean (\(F_2\)) is obtained considering the genetic structure of the crossed population. At the one locus level, the \(F_2\) generation is in Hardy–Weinberg equilibrium and its gene frequency is the average gene frequency of the two parental populations:

\[
p' = (p + r)/2
\]
The array of genotypic values is

\[ p'^2a : 2p'q'd : q'^2(-a) \]

so that the crossed population mean is

\[ \bar{F}_2 = (p' - q')a + 2p'q'd \]
\[ = \left(\frac{1}{2}\right)(p - q + r - s)a + (p + r)(q + s)d \]

which equals \( \bar{MP} \) and \( \bar{F}_1 \) under the assumption of no dominance. Thus, the excess of \( \bar{F}_2 \) over \( \bar{MP} \) is

\[ [(\frac{1}{2})(p + r)(q + s) - pq + rs)d \]

which equals to

\[ (\frac{1}{2})(ps + rq - pq - rs)d = (\frac{1}{2})(p - r)^2d = (\frac{1}{2})H \]

Therefore, \( \bar{F}_2 \) is greater than \( \bar{MP} \) by an amount that is half the excess expressed in the first cross generation, or in other words, half the heterotic effect is lost from the first to the second generation under the restriction of no epistasis when several loci are considered. The complete heterotic effect would be obtained by summation of effects over all loci. In notation of Falconer and Mackay (1996) this becomes

\[ F_2 = \sum (\frac{1}{2})y^2d \]

When considering the cross between two inbred lines, the only difference is that for one line \( p = 0 \) or \( p = 1 \), depending on whether the gene is in homozygous recessive or homozygous dominant condition. For the other line, \( r = 0 \) or \( r = 1 \) for the same locus. The heterotic response in the first-generation cross is due to the loci where \( p = 1 \) and \( r = 0 \) or vice versa, so the heterotic effect depends on the number of such contrasting loci and also on the level of dominance at each locus. In any case, heterotic response is expected to occur whenever there is difference in gene frequencies and some degree of directional dominance at one or more loci involved in the control of the character.

### 10.5 Heterosis and Prediction Methods Across Genotypes

Prediction of means, as well as prediction of results from selection, is one of the important contributions of quantitative genetics to plant breeding (Hallauer, 2006). One of the contributions of genetics to agriculture experimentation has been predicting results following controlled crosses. Following the rediscovery of Mendelian laws, prediction methods have been used extensively in the study of qualitative
traits. In quantitative traits, parameters, such as means rather than proportions of genotypes, are of greater importance. Theory developed for prediction purposes has brought about a better understanding of the nature of gene action and its relation to population means and their components. In the following sections several cases are considered where prediction of population means has possible uses. The theory involved is restricted to the specific conditions of parental populations that have diploid segregation, no preferential fertilization, Hardy–Weinberg equilibrium, linkage equilibrium, and/or negligible epistasis. The procedures presented are useful when a large number of possible crosses cannot be produced and evaluated in experimental trials.

10.5.1 Populations Under Selfing

Theories about prediction involving homozygous (pure) lines were first developed for self-fertilizing crops, but they are useful in maize and other cross-pollinated crops where inbred lines are used (Hallauer, 2006). The following formulas are given by Mather (1949) and Mather and Jinks (1971). Let $P_1$ and $P_2$ represent two completely homozygous lines and $F_1$ the first-generation cross between them. For any quantitative trait the mean of the $F_2$ generation (obtained by selfing or sib-mating $F_1$ plants) can be predicted by

$$F_2 = \left(\frac{1}{4}\right)(P_1 + P_2 + 2F_1)$$

where $P_1$ and $P_2$ are the parent line means and $F_1$ the first-generation cross mean. In the same manner other advanced populations means ($F_3, F_4,$ etc.) may be predicted by

$$F_3 = \left(\frac{1}{8}\right)(3P_1 + 3P_2 + 2F_1), F_4 = \left(\frac{1}{16}\right)(7P_1 + 7P_2 + 2F_1), \ldots$$

The general formula to predict population means in any generation under selfing without selection is

$$F_n = \left(\frac{1}{2}\right)[1 - \left(\frac{1}{2}\right)^{n-1}](P_1 + P_2) + \left(\frac{1}{2}\right)^{n-1}F_1$$

where the $n$th generation is obtained after $n - 1$ generations under selfing (Mather 1949). When the $F_1$ is backcrossed to either parent (BC1 and BC2), its mean is predicted by

$$BC1 = \left(\frac{1}{2}\right)(P_1 + F_1), BC2 = \left(\frac{1}{2}\right)(P_2 + F_1)$$

The $F_2$ mean also is expected to be

$$F_2 = \left(\frac{1}{2}\right)(BC1 + BC2)$$
10.5.2 Double Crosses and Three-Way Crosses

Single-cross hybrids have replaced double-cross hybrids in maize production. However, double-cross hybrids have been extensively used in maize since their use was suggested by Jones (1918) as a solution for hybrid seed production when inbred lines were poor seed producers. In subsequent years, maize breeders’ primary objective has been the development of new superior hybrids. Double-cross hybrids resulted from crosses between two single crosses that were themselves the result of crosses between two inbred lines. The best results were expected to occur when four different inbred lines were used. If the same inbred line was included in both parental single crosses some inbreeding was expected in the double cross, thus precluding the maximization of heterosis. The theory follows.

A set of \( n \) inbred lines can be subdivided into \( C_n^4 \) subsets of four inbred lines. Within each subset only three distinct double crosses are possible. Within the subset \([A, B, C, D]\), the following double crosses are possible: \((A \times B) \times (C \times D)\), \((A \times C) \times (B \times D)\), and \((A \times D) \times (B \times C)\). Thus the total number of possible double crosses is

\[
N_{dc} = 3 C_n^4 = \binom{1}{3} \binom{n(n-1)(n-2)(n-3)}{1}
\]

For example, for \( n = 10 \), then 630 distinct double crosses are possible. Such a number of entries were difficult to evaluate in experimental trials at the time of double-cross hybrid seed production and the need of predictions was evident. Prediction based on single crosses \( (N_{sc} = C_n^2) \) would require only 45 single crosses to get the information for 630 double crosses. Double-cross hybrid breeding programs often used two genetically divergent populations. Theory and empirical data have demonstrated that superior hybrids are more likely to be obtained if both populations are improved by some method of recurrent selection, especially reciprocal recurrent selection (Hallauer, 1973; Suwantaradon and Eberhart 1974; Moll et al. 1977; Eyherabide and Hallauer, 1991; Keeratinijakal and Lamkey, 1993). In such instances the best procedure is to develop one parental single cross from each base population to maximize heterotic effects in the double crosses. If \( n_1 \) inbred lines are available from one population and \( n_2 \) from another, the number of possible double crosses is

\[
N_{dc} = C_{n_1}^2 C_{n_2}^2
\]

In this case only \( n_1 n_2 \) single crosses are needed for prediction. If \( n_1 = n_2 = n \), then

\[
N_{dc} = \binom{1}{3} \binom{n(n-1)}{1}^2
\]

and \( n^2 \) single crosses are needed to predict double crosses.

Three-way crosses have also been successfully used in commercial maize hybrids. They result from crosses between one single cross and one inbred line with the single cross used as the female parent for seed production. Prediction of
three-way crosses also was important whenever the number of inbred lines was too
great for experimental evaluation.
If a fixed set of \( n \) inbred lines is available, the number of possible three-way
crosses is

\[
N_{tc} = 3C_n^3 = \binom{n}{3} = \frac{n(n-1)(n-2)}{6}
\]

For example, for \( n = 10 \) we have \( N_{tc} = 360 \).
If two sets of \( n_1 \) and \( n_2 \) inbred lines are from two distinct populations, we have

\[
N_{tc} = n_1C_{n_2}^2 = n_1n_2(n_2 - 1)/2 \text{ when the single cross is from set 2}
\]

\[
N_{tc} = n_2C_{n_1}^2 = n_1n_2(n_1 - 1)/2 \text{ when the single cross is from set 1}
\]

Total number of possible three-way crosses is

\[
\binom{n_1}{2}n_2(n_1 + n_2 - 2)
\]

In this case only \( n_1n_2 \) single crosses are needed for prediction. If \( n_1 = n_2 = n \), then \( n^2 \) single crosses are needed. For example, if \( n_1 = n_2 = 5 \), then

\[
N_{tc} = \binom{5}{2}25(8) = 100
\]

Prediction of double-cross performance in maize was first reported by Jenkins (1934) using single-cross data. Jenkins suggested four alternative methods for prediction:

A. Mean performance of six possible single crosses among any set of four inbred lines
B. Average performance of four non-parental single crosses
C. Average performance of four inbred lines over a series of single crosses
D. Average performance of a set of four lines when tested by the testcross procedure.

The four methods of prediction differ with respect to type of gene action involved. Methods A, C, and D are related only to additive gene action, while method B involves additive as well as non-additive (dominance and various types of epistasis) effects. Jenkins, using all methods, found a significant correlation between observed and predicted means. However, the correlation was greatest in method B, which agrees with quantitative genetics theory. The efficiency of method B also was reported by several authors (Doxtator and Johnson, 1936; Anderson, 1938; Hayes et al., 1943, 1946).

Prediction of double-cross hybrids using Jenkins’s method B is as follows. In each set of four inbred lines (say \( P_1, P_2, P_3, \) and \( P_4 \)), the six possible single crosses are \( S_{12}, S_{13}, S_{14}, S_{23}, S_{24}, \) and \( S_{34} \). The three possible double crosses may be predicted as
\[ S_{12} \times S_{34} : D_{12.34} = \left( \frac{1}{4} \right)(S_{13} + S_{14} + S_{23} + S_{24}) \]
\[ S_{13} \times S_{24} : D_{13.24} = \left( \frac{1}{4} \right)(S_{12} + S_{14} + S_{23} + S_{24}) \]
\[ S_{14} \times S_{23} : D_{14.23} = \left( \frac{1}{4} \right)(S_{12} + S_{13} + S_{24} + S_{34}) \]

A simple model can be used to illustrate the theory of the prediction procedure, although a rather complex theory may be involved. Suppose that the four parental lines have the following genotypes: P1:AABB, P2:AAbb, P3:aaBB, and P4:aabb. We identify \( a \) (or \(-a\)) and \( d \) as the genotypic effects for homozygotes and heterozygotes and are defined as deviations around a mean \( \mu \) of the two extreme homozygotes. The following effects can be assigned in a diallel table, not considering epistatic effects, where

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
\mu + a_A + a_B & \mu + a_A + a_B & \mu + d_A + a_B & \mu + d_A + a_B \\
\mu + a_A - a_B & \mu + d_A + a_B & \mu + d_A - a_B & \mu - a_A + d_B \\
\mu - a_A + a_B & \mu + d_A - a_B & \mu - a_A + d_B & \mu - a_A - a_B \\
\end{array}
\]

Double cross \( D_{12.34} \) results from the following cross:

<table>
<thead>
<tr>
<th>Parental Single crosses</th>
<th>Double cross Genotypes</th>
<th>Genotypic effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>AABb</td>
<td>AaBB</td>
<td>( \mu + d_A + a_B )</td>
</tr>
<tr>
<td>aaBb</td>
<td>AaBb</td>
<td>( \mu + d_A + d_B )</td>
</tr>
<tr>
<td>aaBb</td>
<td>AaBb</td>
<td>( \mu + a_A + d_B )</td>
</tr>
<tr>
<td>Aabb</td>
<td>Aabb</td>
<td>( \mu + d_A - a_B )</td>
</tr>
<tr>
<td>Average</td>
<td>Average</td>
<td>( \mu + d_A + (\frac{1}{2})d_B )</td>
</tr>
</tbody>
</table>

Average genotypic effects of a double cross can be predicted by average genotypic effects of the four genotypes in the upper right-hand corner of the diallel table. Those effects are relative to the non-parental single crosses. Following the same procedure, three-way crosses of a set of three inbred lines can be predicted by

\[
\begin{align*}
T_{12.3} &= \left( \frac{\sqrt{2}}{2} \right)(S_{13} + S_{23}) \\
T_{13.2} &= \left( \frac{\sqrt{2}}{2} \right)(S_{12} + S_{23}) \\
T_{23.1} &= \left( \frac{\sqrt{2}}{2} \right)(S_{12} + S_{13})
\end{align*}
\]

Eberhart (1964) presented the following formulas to predict double crosses:

1. \( \hat{D}_{ij}^{sa} = \left( \frac{1}{6} \right)(S_{ij} + S_{ik} + S_{il} + S_{jk} + S_{jl} + S_{kl}) \)
2. \( \hat{D}_{ij}^{sb} = \left( \frac{1}{4} \right)(S_{ik} + S_{il} + S_{jk} + S_{jl}) \)
3. \( \hat{D}_{ij}^{ss} = \left( \frac{1}{2} \right)(T_{ij,k} + T_{ij,l}) \)
4. \( \hat{D}_{ij,kl} = (\frac{1}{2})(T_{kl,i} + T_{kl,j}) \)

5. \( \hat{D}_{ij,kl} = (\frac{1}{2})(D_{ij,kl}^{ii} + D_{ij,kl}^{ik,l}) \)

The first two formulas correspond to methods A and B proposed by Jenkins (1934). The others are based on three-way cross performance and have not been extensively used because from a fixed set of inbred lines there are more possible three-way crosses than single crosses. Such formulas may be useful when a desirable single cross \((S_{ij})\) is available and two new inbred lines \((k\) and \(l\)) must be developed to form the double cross \(D_{ij,kl}\).

All of the five formulas mentioned above are unbiased by additive effects and dominance effects, except formula 1 that is unbiased by only additive effects. Hence, formulas 2, 3, 4, and 5 can be efficiently used to predict double crosses if epistatic effects are negligible. When epistatic effects of the dominance type are unimportant relative to other types of epistasis, the following linear relation is suggested to predict the double-cross hybrids (Eberhart, 1964):

6. \( \hat{D}_{1234}^{t-s} = 2D_{1234}^{t} - D_{1234}^{vb} \)

\[= (\frac{1}{2})(T_{12,3} + T_{12,4} + T_{34,1} + T_{34,2}) - (\frac{1}{4})(S_{13} + S_{14} + S_{23} + S_{24}) \]

Eberhart et al. (1964) used formulas 2, 5, and 6 to predict double-cross performance. Although observed values were not available for comparisons, they concluded that differences in predicted values were not significant among the methods. In addition, although epistasis was present it was not of sufficient magnitude relative to experimental error and genotype by environment interaction. Therefore, it did not provide an advantage to formula 6.

The joint number of single crosses and three-way crosses required to predict double crosses is

\[n_s + n_t = (\frac{1}{2})n(n - 1)^2\]

which only exceeds the number of possible double crosses when the number of inbred lines \(n = 7\). If two distinct populations are used, then

\[n_s + n_t = (\frac{1}{2})n_1n_2(n_1 + n_2)\]

where \(n_1\) and \(n_2\) are the number of inbred lines from each population. So

\[n_d = (n_s + n_t) \text{ when } n_1n_2 - 3n_1 - 3n_2 + 1 = 0\]

If \(n_1 = n_2\), then \(n_d > (n_s + n_t)\) when \(n > 5\). Prediction based on single-and three-way crosses is justified only when \(n_d\) is large. But \(n_s + n_t\) is also large as shown in Table 10.7. Hence prediction based on this procedure has a practical limitation.
Table 10.7 Number of possible double crosses \((n_d)\) and number of entries \((n_s + n_t)\) needed for their prediction

<table>
<thead>
<tr>
<th>(n)</th>
<th>(n_d)</th>
<th>(n_s)</th>
<th>(n_s + n_t)</th>
<th>(n')</th>
<th>(n_d)</th>
<th>(n_s + n_t = (n')^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>6</td>
<td>18</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>10</td>
<td>40</td>
<td>3</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>15</td>
<td>75</td>
<td>4</td>
<td>36</td>
<td>64</td>
</tr>
<tr>
<td>7</td>
<td>105</td>
<td>21</td>
<td>126</td>
<td>5</td>
<td>100</td>
<td>125</td>
</tr>
<tr>
<td>8</td>
<td>210</td>
<td>28</td>
<td>196</td>
<td>6</td>
<td>225</td>
<td>216</td>
</tr>
<tr>
<td>9</td>
<td>378</td>
<td>36</td>
<td>288</td>
<td>7</td>
<td>441</td>
<td>343</td>
</tr>
<tr>
<td>10</td>
<td>630</td>
<td>45</td>
<td>405</td>
<td>8</td>
<td>784</td>
<td>512</td>
</tr>
<tr>
<td>15</td>
<td>4,095</td>
<td>105</td>
<td>1,470</td>
<td>9</td>
<td>1,296</td>
<td>729</td>
</tr>
<tr>
<td>20</td>
<td>14,435</td>
<td>190</td>
<td>3,610</td>
<td>10</td>
<td>2,025</td>
<td>1,000</td>
</tr>
<tr>
<td>30</td>
<td>82,215</td>
<td>435</td>
<td>12,615</td>
<td>20</td>
<td>36,100</td>
<td>8,000</td>
</tr>
</tbody>
</table>

Cockerham (1967) presented a unified theory that makes use of both the genetic and experimental conditions to predict double crosses from single-cross hybrid data. The random sample approach was considered and the optimum predictor was suggested. Cockerham found that differences in efficiency when comparing the optimum predictor with Jenkins’s methods A and B were small, although the optimum predictor was generally more efficient. Otsuka et al. (1972) used the concept of optimum predictor to develop equations for use in double-cross, three-way cross, and single-cross estimation, based on general and specific effects estimated from a diallel analysis of fixed effects. The models were

\[
\hat{S}_{ij} = m + \hat{g}_i + \hat{g}_j + \hat{s}_{ij}
\]

\[
\hat{D}_{ij,kl} = m + (\frac{1}{2})(\hat{g}_i + \hat{g}_j + \hat{g}_k + \hat{g}_l) + \lambda(\frac{1}{4})(\hat{s}_{ik} + \hat{s}_{il} + \hat{s}_{jk} + \hat{s}_{jl})
\]

\[
\hat{T}_{ij,k} = m + (\frac{1}{2})(\hat{g}_i + \hat{g}_j) + \hat{g}_k + \lambda(\frac{1}{2})(\hat{s}_{ik} + \hat{s}_{jk})
\]

where \(\lambda\), varying between 0 and 1, is a weighting coefficient for specific effects.

When \(\lambda = 1\) the formula for double-cross prediction corresponds to Jenkins’s method B. The use of optimum weight predictor was possible, but it was found that prediction with Jenkins’s method B \((\lambda = 1)\) was nearly as efficient. The authors also found that Jenkins’s method A was slightly superior to any other predictor in some instances, suggesting that information from parental single crosses may be effective for predicting hybrids from a fixed set of highly selected lines.

Accuracy of prediction depends more on number of replications and environments than on small differences in prediction methods. Otsuka et al. (1972) noted, for example, that 45 single crosses from 10 lines would require 450 field plots (two replications and five environments) to predict 630 possible double crosses. On the other hand, 2,520 plots (two locations and two replications) would have been required to have a similar precision on double-cross performance itself.
10.5.3 Synthetic and Composite Varieties

Synthetic varieties have been widely used for commercial and breeding purposes since they were first suggested by Hayes and Garber (1919). They were defined by Lonnquist (1961) as ‘open-pollinated populations derived from the intercrossing of selfed plants or lines and subsequently maintained by routine mass selection procedures from isolated plantings.’ When open-pollinated varieties instead of inbred lines are intercrossed, resulting populations are usually called composites or composite varieties. Synthetic and composite varieties are very similar in structure, but the distinction may be useful for practical purposes. Some breeders use synthetic variety in its broadest sense, including any open-pollinated population derived from artificial selection; to avoid confusion, the term is used herein according to Hayes and Garber’s (1919) concept (defined by Lonnquist, 1961).

Yield or any quantitative traits of synthetic varieties are predicted according to a formula based on Wright’s (1922) statement: ‘A random-bred stock derived from $n$ inbred families will have $(1/n)\text{th}$ less superiority over its inbred ancestry than the first cross or a random-bred stock from which the inbred families might have been derived without selection.’ The formula developed from this principle is commonly called Wright’s formula and was cited by Kinman and Sprague (1945) as

$$\hat{Y}_2 = \hat{Y}_1 - (\hat{Y}_1 - \hat{Y}_0)/n$$

where

- $\hat{Y}_2$ = the mean of a synthetic variety obtained by intercrossing all possible single crosses among a set of $n$ inbred lines,
- $\hat{Y}_1$ = average performance of all single crosses among $n$ inbred lines, and
- $\hat{Y}_0$ = average performance of $n$ parental inbred lines.

The closeness of observed and predicted means using Wright’s formula has been reported (e.g., Neal, 1935).

A general formula to predict synthetic varieties was given by Busbice (1970):

$$\overline{Y}_t = \overline{Y}_0 + [(F_0 - F_t)/(F_0 - F_1)](\overline{Y}_1 - \overline{Y}_0) \quad F_0 \neq F_1$$

where $F_t (i = 0, 1, 2, \ldots, t)$ is Wright’s coefficient of inbreeding in the $i$th generation. Values of $F_1$ and $F_t$ can be computed from $F_0$, using special formulas where the coefficient of parentage of parents ($r_0$), the ploidy number ($2k$), the frequency of selfing ($s$), and the number of parents ($n$) are considered. Terms $\overline{Y}_0$, $\overline{Y}_1$, and $\overline{Y}_t$ are the means in generations 0 (parents), 1 (all possible single crosses among $n$ parents), and $t$. If the parents are completely homozygous ($F_0 = 1$), unrelated, diploid ($2k = 2$), and no selfing occurs ($s = 0$), the general formula reduces to Wright’s formula.

Gilmore (1969) concluded that Wright’s formula can be used when lines are at any level of inbreeding. He showed that Wright’s formula is valid for $S_1$, $S_2$, $\ldots$, $S_n$.
lines; parental lines need not be limited to those developed from selfing, and gene frequency must be limited to 1 or 0.5. It is only required that parental lines be in Hardy–Weinberg equilibrium for each locus. As pointed out by Mochizuki (1970), Wright’s formula is used to predict composite populations because parental varieties are in genetic equilibrium.

### 10.5.4 Composite Populations

Composite populations result from random mating all possible inter-varietal crosses among a fixed set of heterogeneous varieties (e.g., populations or races). Composites have been widely used as breeding populations because greater genetic variability is expected to be available if populations of diverse origins are combined. For instance, diverse populations can include exotic germplasm. Heterosis among inter-varietal crosses also has been found to be relatively high. Therefore, the population-hybrid concept (Carena and Wicks III, 2006) follows the inbred–hybrid concept. When a composite is formed the mean yield of the new population is expected to be greater than the average of the parental varieties.

Selection within and among parental varieties may give a higher composite mean. In some instances poorer varieties can be discarded and the breeder can start the program at a higher level of productivity. Choosing varieties to be used as parents may be based on a variety cross diallel evaluated in a series of environments (Tables 10.3, 10.4, 10.5). Composite means thus can be predicted from the diallel data.

The number of possible different composites $N_{co}$ increases greatly with an increase in the number of parental varieties (Vencovsky and Miranda, 1972):

$$N_{co} = 2^n - (n + 1)$$

For $n = 10$ parental varieties, it is possible to have 1,013 different composites for the cases in which each parental variety contributes equal proportions of germplasm. Actually the number of possible composites is infinite if we include those types where unequal contributions of parental varieties is assumed. The formula above also gives the number of possible synthetic varieties from a set of $n$ inbred lines. Prediction of composite means was first suggested by Eberhart et al. (1967), where an expression was similar to Wright’s formula:

$$\hat{Y}_{co} = \bar{Y}_c - (\bar{Y}_c - \bar{Y}_v)/n$$

where:

$\hat{Y}_{co}$ = the predicted mean of a quantitative trait for a composite obtained from random mating,

$\bar{Y}_c$ = the average of all possible inter-varietal crosses among $n$ parental varieties,

$\bar{Y}_v$ = the mean of $n$ parental varieties, and

$n =$ the number of varieties.
The predicted means are obtained from a diallel table as follows. Taking the complete diallel table as reference, each predicted mean is obtained by averaging all \( n^2 \) values of a partial diallel table containing only the values from parental varieties and their crosses relative to the composite under consideration. An example of a reference diallel table is shown below from which data are used to predict the composite means \( A (\bar{Y}_{123}) \) and \( B (\bar{Y}_{456}) \):

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>71</td>
<td>67</td>
<td>69</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>58</td>
<td>62</td>
<td>62</td>
<td>58</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>46</td>
<td>60</td>
<td>55</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>55</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>47</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Composite A: (123)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>71</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>58</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>62</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>

Mean: \( Y_{123} = 62.7 \)

Composite B: (456)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>55</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>47</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>51</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>

Mean: \( Y_{456} = 50.7 \)

Mochizuki (1970) showed a detailed derivation of the above formula, presenting also its alternative forms as follows:

\[
Y_{co}^r = \mu + (1/n)(a_1 + a_2 + \cdots + a_n) + (1/n)(d_1 + d_2 + \cdots + d_n) \\
+ (2/n^2)(h_{12} + h_{13} + \cdots + h_{n-1,n})
\]

\[
Y_{co}^r = \mu + (1/n) \sum_{j=1}^{n} (a_j + d_j) + (2/n^2) \sum_{j<j'} h_{jj'}
\]

where \( \mu \) is the parent varieties mean; \( a_j \) and \( d_j \) are deviations due to homozygotes and heterozygotes relative to the \( j \)th variety; and \( h_{jj'} \) is the total heterosis relative to cross between varieties \( j \) and \( j' \).

In terms of observed means it follows that

\[
Y_{co}^r = \frac{2}{n(n-1)} \sum_{j<j'} Y_{jj'} - \frac{1}{n} \left[ \frac{2}{n(n-1)} \sum_{j<j'} Y_{jj'} - \frac{1}{n} \sum_j Y_{jj'} \right]
\]

\[
= \bar{Y}_c - \frac{\bar{Y}_c - \bar{Y}_v}{n}
\]
When the objective of a recurrent selection program is hybrid development, the recommended procedure is the use of two base populations with some genetic divergence between them. Starting from a set of \( n \) varieties, a pair of composites can be formed so that a good complementation, or a good specific combining ability, exists between them. Such a pair of base populations may be efficiently used for a reciprocal recurrent selection or for the development of inbred lines and hybrids. The number of possible pairs of composites is given by (Vencovsky and Miranda, 1972):

\[
N_{pc} = \left(\frac{1}{2}\right)[3^n - 2^n(n + 2) + n(n + 1) + 1]
\]

where varieties in each composite contribute equally to predict each pair of composites.

For instance,

\[
\hat{Y}_{A \times B} = \frac{1}{(mn)}(Y_{11} + Y_{12} + \cdots + Y_{1n} + Y_{21} + Y_{22} + \cdots + Y_{2n} + \cdots + Y_{m1} + Y_{m2} + \cdots Y_{mn})
\]

where \( Y_{ij} \) is a cross of the \( i \)th variety \((i = 1, 2, \ldots, m)\) in composite A and the \( j \)th variety \((j = 1, 2, \ldots, m)\) in composite B.

For example, from the reference diallel table the composite cross \((123) \times (456)\) mean is predicted by

\[
\hat{Y}_{A \times B} = \left(\frac{1}{9}\right)(69 + 63 + \cdots + 54) = 60.0
\]

and the composite cross \((135) \times (246)\) mean is predicted by

\[
\hat{Y}_{A' \times B'} = \left(\frac{1}{9}\right)(71 + 69 + \cdots + 51) = 60.2
\]

as shown in the following example:

<table>
<thead>
<tr>
<th>Composite cross: ((123) \times (456))</th>
<th>Composite cross: ((135) \times (246))</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>71</td>
</tr>
<tr>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>Mean: ( \hat{Y}_{A \times B} = 60 )</td>
<td>Mean: ( \hat{Y}_{A' \times B'} = 60.2 )</td>
</tr>
</tbody>
</table>

Note that the prediction formulas for composites and pairs of composites are valid when parental varieties are in genetic equilibrium and when they contribute equally in each composite. In the case of unequal contribution of parental varieties, a general formula is given by Vencovsky (1970). Denoting by \( p_{ij} \) the frequency of the (favorable) allele at the \( i \)th locus in the \( j \)th variety in the composite population the gene frequency is
\[ \bar{p}_i = \sum_j f_j p_{ij}, \]

where \( f_j \) is the proportion of \( j \)th variety germplasm in the composite and \( \sum f_j = 1 \) (Wahlund, 1928).

At genetic equilibrium the composite mean is

\[ \hat{Y}_{co} = \mu' + \sum_j (2\bar{p}_i - 1)a_i + 2 \sum_i (2\bar{p}_i - \bar{p}_i^2)d_i \]

Introducing parameters defined by Gardner and Eberhart (1966), this formula may be expressed as

\[ \hat{Y}_{co} = \mu' + \sum_j f_j (\hat{a}_j + \hat{d}_j) + 2 \sum_{j < j'} f_j f_{j'} \hat{h}_{jj'} \]

this also can be expressed in terms of observed means as

\[ \hat{Y}_{co} = \sum_j f_j Y_{jj} + 2 \sum_{j < j'} f_j f_{j'} Y_{jj'} \]

where \( Y_{jj} \) and \( Y_{jj'} \) are taken from a diallel table.

In the same way, a cross between two composites can be predicted as

\[ \hat{Y}_{A \times B} = \sum_{jj'} f_j f_{j'} Y_{jj'} \]

where \( j \) and \( j' \) refer to varieties that enter in composites A and B, respectively.

Prediction of composite means has been used by several breeders for both equal contribution and unequal contribution of parental varieties (Hallauer and Eberhart, 1966; Darrah et al., 1972; Vencovsky et al., 1973; Miranda, 1974a). Prediction of pairs of composites as a first step for a reciprocal recurrent selection program also has been used (Vencovsky et al., 1973).

### 10.5.5 Generalization of Prediction Methods

When epistasis is taken as negligible, several formulas are available for prediction of double crosses, three-way crosses, composites, and synthetics. A general formula was suggested by Vencovsky (1973) to predict means in any non-inbred population obtained through controlled crossing if the parents are in genetic equilibrium:

\[ \bar{Y} = \left( \sum_i f_i P_i \right) \left( \sum_j f_j P_j \right) \]
In the above expression, it is assumed that the resulting population comes from two original sources of gametes (female and male gamete sources). Hence the expression within the first parentheses represents a male (or female) gamete source and the second a female (or male) gamete source; \( f_i \) and \( f_j \) represent the proportion of parental germplasm (original) in the resulting population, so \( \Sigma f_i = \Sigma f_j = 1 \). Terms \( P_i \) and \( P_j \) denote symbolically the parental (original) types. After parents have been identified and their proportions assigned, the expression must be algebraically extended and terms \( P_i^2 \) or \( P_j^2 \) replaced by \( Y_i \) or \( Y_j \) (parent means) and cross-products \( P_iP_j \) and \( P_jP_i \) (assuming no differences between reciprocal crosses) by \( Y_{ij} \) (hybrid means).

A numerical example may illustrate the application of the general formula. Let the diallel table below represent means of a quantitative character, where any parental type is in Hardy–Weinberg equilibrium:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
<td>1.8</td>
<td>2.0</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>1.6</td>
<td>1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Three-way cross prediction:

\[
\hat{T}_{12.3} = [(\frac{1}{2})P_1 + (\frac{1}{2})P_2](P_3)
\]

\[
= (\frac{1}{2})P_1P_3 + (\frac{1}{2})P_2P_3
\]

and after transformation,

\[
\hat{T}_{12.3} = (\frac{1}{2})Y_{13} + (\frac{1}{2})Y_{23}
\]

\[
= (\frac{1}{2})(2.0) + (\frac{1}{2})(1.6) = 1.80
\]

A three-way cross like \( \hat{T}_{12.3} \) results from a cross between a single cross (\( S_{12} \)) and an inbred line (\( P_3 \)). If the single cross is used as the female, the original sources of female gametes are the inbred lines \( P_1 \) and \( P_2 \) with equal participation (\( f_1 = f_2 = \frac{1}{2} \)) and are represented in the first term. In the same way the original source of male gametes is the inbred line \( P_3 \), which is represented in the second term.

(b) Double-cross prediction:

\[
\hat{D}_{12.34} = [(\frac{1}{2})P_1 + (\frac{1}{2})P_2][(\frac{1}{2})P_3 + (\frac{1}{2})P_4]
\]

\[
= (\frac{1}{4})(Y_{13} + Y_{14} + Y_{23} + Y_{24}) = 1.78
\]

(c) Prediction of an \( F_2 \) generation from the cross \( S_{12} \):

\[
\bar{F}_2 = [(\frac{1}{2})P_1 + (\frac{1}{2})P_2][(\frac{1}{2})P_1 + (\frac{1}{2})P_2]
\]

\[
= (\frac{1}{4})(\bar{Y}_1 + 2\bar{Y}_2) = 1.65
\]
Note that an $F_2$ generation is obtained either by selfing $F_1$ plants when parents are inbred lines or by crossing among themselves. For both instances, the female and male gametes come from the same original source ($P_1$ and $P_2$).

(d) Prediction of a synthetic variety or a composite with equal contributions. For instance, $f_1 = f_2 = f_3 = f_4 = \frac{1}{4}$:

$$Y_{\text{syn}} = \left[ (\frac{1}{4})P_1 + (\frac{1}{4})P_2 + (\frac{1}{4})P_3 + (\frac{1}{4})P_4 \right]^2$$

$$= (\frac{1}{16})(\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3 + \bar{Y}_4)$$

$$+ \frac{1}{8}(\bar{Y}_{12} + \bar{Y}_{13} + \bar{Y}_{14} + \bar{Y}_{23} + \bar{Y}_{24} + \bar{Y}_{34})$$

$$= 1.74$$

The gametes that form a synthetic variety come from the same original source, which is represented by the parental inbred lines. The same situation holds for the composites.

(e) Prediction of a synthetic variety or a composite with unequal contributions. For instance, $f_1 = \frac{1}{2}, f_2 = \frac{1}{6}, f_3 = \frac{1}{6},$ and $f_4 = \frac{1}{6}$:

$$\bar{Y}_{\text{syn}} = \left[ (\frac{1}{2})P_1 + (\frac{1}{6})P_2 + (\frac{1}{6})P_3 + (\frac{1}{6})P_4 \right]^2$$

$$= (\frac{1}{4})\bar{Y}_1 + (\frac{1}{36})(\bar{Y}_{12} + \bar{Y}_{13} + \bar{Y}_{14}) + \frac{1}{16}(\bar{Y}_{23} + \bar{Y}_{24} + \bar{Y}_{34})$$

$$= 1.74$$

(f) Prediction of a backcross, for instance, $Y_{12.2}$:

$$Y_{12.2} = [(\frac{1}{2})P_1 + (\frac{1}{2})P_2](P_2)$$

$$= (\frac{1}{2})\bar{Y}_{12} + (\frac{1}{2})\bar{Y}_2$$

$$= 1.6$$

10.5.6 Heterosis as a Component of Means

Heterosis or hybrid vigor is regarded as the superiority of a hybrid over its parents. Its quantitative measure usually is in relation to the average of the parents:

$$H = \bar{F}_1 - (\bar{P}_1 + \bar{P}_2)/2 \quad \text{or} \quad \bar{F}_1 = (\bar{P}_1 + \bar{P}_2)/2 + H$$

which is applied for both cross- and self-pollinated species. The mean of any derived population from an $F_1$ cross may be associated with changes in heterosis, e.g., for the $F_2$ generation, obtained by random mating in the $F_1$ generation:

$$\bar{F}_2 = (\bar{P}_1 + \bar{P}_2)/2 + (\frac{1}{2})H$$

This expression also is valid for selfed $F_1$ when parents are homozygous lines. It can be seen that half the heterotic effect is lost after the first generation of selfing. With
continuous selfing, the heterosis in each generation is expected to be half of the one in the preceding generation. So the general formula to predict an \( F_n \) generation is

\[
F_n = (\overline{P}_1 + \overline{P}_2)/2 + (\frac{1}{2})^{n-1}H
\]

where \( F_n \) is a population after \( n - 1 \) generations under selfing (Mather and Jinks, 1971). The decrease in heterosis by half in each generation is explained by the decrease in the frequency of heterozygous loci. For one locus we have

<table>
<thead>
<tr>
<th>Generation</th>
<th>Genotypes</th>
<th>Frequency of heterozygotes, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_1 )</td>
<td>( Aa )</td>
<td>100</td>
</tr>
<tr>
<td>( F_2 )</td>
<td>( \frac{1}{4}AA )</td>
<td>50</td>
</tr>
<tr>
<td>( F_3 )</td>
<td>( \frac{1}{8}AA )</td>
<td>25</td>
</tr>
<tr>
<td>( \cdots )</td>
<td>( \cdots )</td>
<td>( \cdots )</td>
</tr>
<tr>
<td>( F_n )</td>
<td>( (\frac{1}{2})^{n-1}AA )</td>
<td>100((\frac{1}{2}))^{n-1}</td>
</tr>
</tbody>
</table>

When several generations of selfing are evaluated with the parents, several estimates for heterosis are available. When the data have a good fit to the theoretical model, the best estimate for heterosis as well as for parent means is the least squares estimate. A general least squares formula may be useful for estimating heterosis:

\[
H = \frac{(n + 2) \left[ \sum_{k=1}^{n} \left( \frac{1}{2} \right)^k \overline{F}_k \right] - \left( \frac{2^{n-1}}{2^n} \right) G}{2 \left[ \frac{(n+2)}{3} \frac{(4^{n-1})}{4^n} - \left( \frac{2^{n-1} \cdot 3}{2^n} \right) \right]}
\]

where

\[
G = \overline{P}_1 + \overline{P}_2 + \overline{F}_1 + \overline{F}_2 + \cdots + \overline{F}_n \quad \text{and} \quad \sum_{k=1}^{n} \left( \frac{1}{2} \right)^k \overline{F}_k = \left( \frac{1}{2} \right) \overline{F}_1 + \left( \frac{1}{4} \right) \overline{F}_2 + \left( \frac{1}{8} \right) \overline{F}_3 + \cdots
\]

Besides selfing, the original heterotic effect can be modified by other types of matings. For backcrosses we have

\[
\overline{BC}_1 = (3\overline{P}_1 + \overline{P}_2)/4 + (\frac{1}{2})H
\]

In a synthetic variety (or composite), means also can be expressed as functions of the average heterosis of all crosses among \( n \) parent inbred lines (or \( n \) varieties in the case of composites):

\[
Y_{syn} = Y_{co} = \overline{P} + \left[ (n - 1)/n \right] \widehat{h}
\]
Table 10.8  Number of composites, number of pairs of composites, and average heterosis retained from use of \( n \) varieties to synthesize the composites

<table>
<thead>
<tr>
<th>( n )</th>
<th>Number of composites(^a)</th>
<th>Number of composite pairs</th>
<th>Average heterosis retained(^b) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>50.0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0</td>
<td>66.7</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>3</td>
<td>75.0</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>25</td>
<td>80.0</td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>130</td>
<td>83.3</td>
</tr>
<tr>
<td>7</td>
<td>120</td>
<td>546</td>
<td>85.7</td>
</tr>
<tr>
<td>8</td>
<td>247</td>
<td>2,037</td>
<td>87.5</td>
</tr>
<tr>
<td>9</td>
<td>502</td>
<td>7,071</td>
<td>88.9</td>
</tr>
<tr>
<td>10</td>
<td>1,013</td>
<td>23,436</td>
<td>90.0</td>
</tr>
<tr>
<td>20</td>
<td>1,048,555</td>
<td>1,731,858,075</td>
<td>95.0</td>
</tr>
<tr>
<td>30</td>
<td>1,073,741,793</td>
<td>102,928,386,178,606</td>
<td>96.7</td>
</tr>
</tbody>
</table>

\(^a\)For equal contribution of each variety to the composite germplasm

\(^b\)When all \( (n) \) varieties are used to synthesize one composite

where \( \bar{P} \) is the parent mean (inbred lines or varieties). Note that a portion of heterotic effects can be retained in composites. Thus the use of composites may be considered as a method of direct utilization of heterosis expressed in inter-varietal crosses. Table 10.8 shows the number of possible composites and pairs of composites when a given number \( n \) of varieties is available. It also indicates the percentage of average heterosis that is theoretically retained when all varieties are used.

In the synthesis of composites it is not always convenient to have equal contributions of all varieties. Sometimes varieties are included because they have specifically defined traits, such as yield, fast dry down, earliness, grain quality, drought and cold tolerance, plant height, disease resistance, lodging resistance. If some of the traits are more important than others, it may be desired to have some varieties contributing a greater proportion to the composite germplasm. A similar situation was discussed by Miranda and Vencovsky (1973), where a high-yielding variety and seven short plant varieties were available to form a composite. It was desired to introduce genes for shorter plants into the high-yielding variety. Because emphasis was mainly for yield, the first variety contributed 50% and the seven short plant varieties the other 50% of the new composite germplasm. In this situation the possible number of composites is \( 2^n - 1 \), and the new composite can be predicted by Miranda (1974a)

\[
Y_{co}^r = (\frac{1}{2})(\bar{Y}_o + \bar{Y}_v) + (\frac{1}{2})h_o + [(n - 1)/(4n)]h_v
\]

where

\( \bar{Y}_o = \) observed mean of the base population,
\( \bar{Y}_v = \) average for the remaining \( n \) varieties (for the given example, \( n = 7 \)),
\( h_o = \) additive genetic variance of the base population,
\( h_v = \) additive genetic variance of the remaining varieties.
Components of Heterosis in Inter-varietal Diallel Crosses

\( \bar{h}_o \) = average heterosis of all crosses between the base population and other varieties, and
\( \bar{h}_v \) = average heterosis of crosses among the short plant varieties.

It can be seen for this example that the new composite can theoretically retain 50% of \( \bar{h}_o \) and 21.4% of \( \bar{h}_v \). A general formula for situations like this one can be derived by assuming that \( k \) sets of varieties (grouped according to defined characteristics) are available. If each set contributes a proportion \( f_i \) \((i = 1, 2, \ldots, k)\) to the new composite germplasm, each variety of the \( i \)th set enters with a proportion \( f_i/n_i \), where \( n_i \) is the number of varieties in the \( i \)th set. Thus the new composite mean may be predicted by

\[
\hat{Y}_{co} = \sum_{i=1}^{k} f_i \bar{Y}_i + \sum_{i=1}^{k} f_i^2 \left[ \frac{(n_i - 1)}{n_i} \bar{h}_i + 2 \sum_{i < i^{'}} f_i f_i^{'} \bar{h}_{ii}^{'} \right]
\]

where
\( \bar{Y}_i \) = average of all varieties in the \( i \)th set,
\( \bar{h}_i \) = average heterosis of all crosses among varieties of the \( i \)th set, and
\( \bar{h}_{ii}^{'} \) = average heterosis of all crosses between varieties of the \( i \)th set and varieties of the \( i^{'} \)th set (Miranda, 1974b).

### 10.6 Components of Heterosis in Inter-varietal Diallel Crosses

A better understanding of the inter-varietal heterosis expressed as components of means was given by Gardner and Eberhart (1966). They expressed an inter-varietal cross mean by:

\[
Y_{jj^{'}} = \mu + (v_j + v_j^{'}) + \bar{h} + h_j + h_j^{' + s_{jj^{'}}}
\]

where
\( \mu \) = mean of \( n \) parental varieties,
\( v_j \) = variety effect, where the \( j \)th variety enters as one of the parents \((j = 1, 2, \ldots, n)\),
\( \bar{h} \) = average heterosis of all crosses,
\( h_j \) = variety heterosis, which is a constant contribution of the \( j \)th variety to heterosis of all crosses where it enters as one of the parents, and
\( S_{jj^{'}} \) = specific heterosis for the cross between varieties \( j \) and \( j^{'} \) and is a deviation from the expected mean based on \( \bar{h} \) and \( h_j \) effects, i.e., \( S_{jj^{'}} = h_{jj^{'}} - \bar{h} - h_j - h_j^{'} \).

All parameters are estimated from an inter-varietal diallel. The Gardner-Eberhart (1966) complete model partitions the total heterosis \( h_{jj^{'}} \) into three components:

\[
h_{jj^{'}} = \bar{h} + h_j + h_j^{'} + s_{jj^{'}}
\]
Other reduced models can be used, where some effects are omitted. Reduced models were suggested by Eberhart and Gardner (1966) when some effects are found to be unimportant from the analysis of variance. In such instances reduced models give a greater precision in the estimates. The models available are

Model 1: \( Y_{jj'} = \mu + \frac{1}{2}(v_j + v_{j'}) \)

Model 2: \( Y_{jj'} = \mu + \frac{1}{2}(v_j + v_{j'}) + \bar{h} \)

Model 3: \( Y_{jj'} = \mu + \frac{1}{2}(v_j + v_{j'}) + \bar{h} + h_j + h_{j'} \)

Model 4: \( Y_{jj'} = \mu + \frac{1}{2}(v_j + v_{j'}) + \bar{h} + h_j + h_{j'} + s_{jj'} \)

Least squares estimates of the parameters included in the models are

<table>
<thead>
<tr>
<th>Parameter estimate</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\mu} )</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>( \hat{v}_j )</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>( \hat{h} )</td>
<td>2, 3, 4</td>
</tr>
<tr>
<td>( \hat{h}_j )</td>
<td>3, 4</td>
</tr>
<tr>
<td>( \hat{s}_{jj'} )</td>
<td>4</td>
</tr>
</tbody>
</table>

Eberhart and Gardner (1966) presented a general model for genetic effects, where heterosis of inter-varietal crosses was defined according to gene frequencies and dominance effects. Thus, if \( p_{ij} \) and \( q_{ij} \) are frequencies of two alleles at the \( i \)th locus in variety \( j \) and \( p_{ij} \) and \( q_{ij} \) are the frequencies of the same alleles in variety \( j' \), the inter-varietal heterosis is defined by

\[
\begin{align*}
    h_{jj'} &= \sum_i (p_{ij} - \bar{p}_i)(q_{ij} - \bar{q}_i)\delta_i = \sum_i (p_{ij} - p_{ij'})^2\delta_i \\
    \text{this can be extended for multiple alleles. An alternative expression was given by Casas and Wellhausen (1968) for diallel crosses as follows:}
    h_{jj'} &= z_j + z_{j'} - 2w_{jj'}
\end{align*}
\]

where

\[
\begin{align*}
    z_j &= \sum_i (p_{ji} - \bar{p}_i)^2d_i \\
    z_{j'} &= \sum_i (p_{j'i} - \bar{p}_{j'i})^2d_i \\
    w_{jj'} &= \sum_i (p_{ji} - \bar{p}_j)(p_{j'i} - \bar{p}_{j'i})d_i
\end{align*}
\]

Vencovsky (1970) used this model (based on \( z_j \) and \( w_{jj'} \)) to permit genetic interpretation of \( \bar{h}, h_j, \) and \( s_{jj'} \). The author found the least squares estimates of \( z_j \) and \( w_{jj'} \) to be
\[ \hat{z}_j = (1/n)\{Y_{ij} - [(n-2)/2]Y_{jj}\} - (1/n^2)(Y_H + Y_v/2) \]

\[ \hat{w}_{jj'} = (1/n^2)(Y_{jj} + Y_{jj'} + Y_{ij} + Y_{ij'}) - (1/n^2)(Y_H + Y_v/2) - \left(\frac{1}{2}\right)Y_{jj'} \]

A first conclusion from these expressions is that if a given variety \( k \) has \( \hat{z}_k = 0 \) (its gene frequency has no deviation from the average), it necessarily has \( \hat{w}_{jk} = 0 \), and consequently

\[ \hat{h}_{jk} = z_j(j = 1, 2, \ldots, n; j \neq k) \]

The term \( z_j \) is the expected heterosis in a cross where variety \( j \) is one of the parents and the other parent is a pool of all other varieties that form the complete diallel.

Components of total heterosis, according to Gardner and Eberhart’s (1966) model, were then demonstrated to be as follows:

(1) **Average heterosis**:

\[ \bar{h} = 2[n/(n-1)]\bar{z} = 2[n/(n-1)]\sum_i \hat{\sigma}_{ip}^2 d_i \]

where

\[ \bar{z} = (1/n)\sum_j z_j \]

and \( \hat{\sigma}_{ip}^2 \) is the variance of gene frequency (locus \( i \)) over varieties.

Thus, average heterosis is zero when \( d_i = 0 \) for all loci or when \( \hat{\sigma}_{ip}^2 = 0 \) for all loci (e.g., no difference in gene frequency among varieties).

When considering only two varieties, \( \bar{h} \) is the heterosis in the variety cross. A recurrent selection program will lead to an increase in heterosis if there is an increase in the variance of gene frequencies. However, if the increase in gene frequency is in the same direction and with the same magnitude in both populations, heterosis will be unchanged.

(2) **Variety heterosis**:

\[ h_j = [n/(n-2)](z_j - \bar{z}) = [n/(n-2)]\left[ \sum_i (\bar{p}_{ji} - \bar{p}_j)^2 d_i - \sum_i \hat{\sigma}_{ip}^2 d_i \right] \]

Basically, \( h_j \) is a deviation of \( z_j \) around \( \bar{z} \). The most negative value among the \( h_j \) will occur when \( z_j = 0 \); i.e., \( p_{ji} = \bar{p}_j \) (meaning that gene frequency equals the average gene frequency of all parents, assuming a positive dominance). Positive values for \( h_j \) will occur in varieties having \( z_j > \bar{z} \). Considering many loci, this situation may occur for

1. varieties with many loci at a high gene frequency,
2. varieties with many loci at a low gene frequency, and
3. varieties that show a dispersion of gene frequency (high and low) relative to the average gene frequency at each locus. Such relations are shown hypothetically in Table 10.9.
Table 10.9  Values for $z_j$ in three varieties that show different gene frequency. Example with four loci with $d_i = 1$ and $\bar{p}_i = 0.6$ (Vencovsky, 1970)

<table>
<thead>
<tr>
<th>Variety</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>$z_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.16</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>0.4</td>
<td>0.8</td>
<td>0.4</td>
<td>0.16</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$\bar{p}_i$</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td></td>
</tr>
</tbody>
</table>

(3) Specific heterosis:

$$s_{jj'} = \frac{2}{(n-2)}[[n/(n-1)]\bar{z}_j - z_j - z_j'] - 2w_{jj'}$$

Specific combining ability or specific heterosis $s_{jj'}$ depends on the size $n$ of the diallel set, on the average heterosis, and on the heterotic component of general combining ability in addition to the $w_{jj'}$ component. For large $n$, $s_{jj'} = -2w_{jj'}$, showing that greater values of specific combining ability are expected for varieties that are more divergent for genes showing dominance effects. Even for large $n$, $s_{jj'}$ is not a fixed property of a specific cross, because it depends also on the average gene frequency $\bar{p}_i$ (Vencovsky, 1970).

The general combining ability effect is given by

$$g_j = \left(\frac{1}{2}\right)(a_j + d_j) + [n/(n-2)](z_j - \bar{z})$$

where $a_j$ and $d_j$ are defined by Gardner and Eberhart (1966) as being related to homozygous and heterozygous contributions, respectively. Eberhart and Gardner (1966) also defined the variety effect as

$$v_j = a_j + d_j'$$

and, the relation for general combining ability as

$$g_j = \left(\frac{1}{2}\right)v_j + h_j$$

The general combining ability effect depends not only on the variety effect but also on dominance effects that arise in inter-varietal crosses, i.e., on the variety heterosis $h_j$. 

10.7 Conclusions

The earlier studies that examined the heterosis expressed in crosses of plants were primarily interested in the contributions of the parents to the crosses and the recovery of parental traits in the selfing generations of the crosses (Zirkle, 1952; Goldman, 1999). They observed that the crosses were usually more vigorous and productive than the parents, but they did not envision the commercial potential of heterosis. Because of the successful commercialization of hybrid maize during the 20th century, interest in the potential of hybrids (and heterosis) has received greater interest in other plant species (Coors and Pandey, 1999). The relative heterosis expressed in crosses depends on the method of calculation and the relative values of the parents crosses, including the following:

(1) Mid-parent (MP) heterosis \(= \frac{(F_1 - MP)}{MP} \times 100 = \% \text{ MPH} \);
(2) High-parent (HP) heterosis \(= \frac{(F_1 - HP)}{HP} \times 100 = \% \text{ HPH} \); and
(3) Absolute heterosis \(= F_1 - MP = \Sigma y^2 d \).

If \(P_1 = 3.58\), \(P_2 = 5.22\), and \(F_1 = 6.44 \text{ t ha}^{-1}\), the estimates of MP, HP, absolute heterosis would be 46.3, 23.4\%, and 2.04 \text{ t ha}^{-1}, respectively.

The relative heterosis can change, depending on the relative values of the parents and hybrids (Duvick, 1999). Estimates of mid-parent heterosis are of interest genetically, but estimates of high-parent heterosis are of greater interest commercially. The expression \(\Sigma y^2 d\) includes the elements necessary for the expression and relative magnitude of heterosis. Therefore, if either element is zero, heterosis is not expressed. Some level of dominance (and epistasis) is necessary to have heterosis and the difference in gene frequency (\(y^2\)) determines the relative magnitude of heterosis. Although extensive and intensive efforts have been done, definitive evidence for the genetic basis of heterosis remains unclear (Coors and Pandey, 1999) even though it was predicted to be explained by molecular approaches. Molecular markers have been effective to profile inbred lines but they have neither been accurate to assign lines to appropriate heterotic groups nor predict the final hybrid. The ultimate selections depend on replicated field trials to determine which cross has the combination of genetic effects that have consistent, high performance under the environmental effects where tested (Hallauer and Carena, 2009). Extensive testing is conducted to identify those crosses that have the unique combination of genetic effects and allele frequencies for the traits of interest.

Hallauer and Carena (2009) emphasized that each hybrid has a unique assemblage of genes and interactions, and the exact genetic basis of each hybrid also will be unique. Some general information (e.g., level of dominance and dominance types epistasis and differences in allele frequencies) can be determined but each hybrid will be different (i.e., specific combining ability). The level of expression is not predictable because of allele interactions, similar to the estimates of SCA. Each single-cross hybrid is a unique cross between two elite inbred lines, and the cross represents the additive genetic effects of each parent, additive × additive epistatic
effects of each parent, interaction effects of the alleles of both parents and the epistatic effects that include dominance, which suggests a very complex genetic system. And the relative importance of each type of genetic effects also will be different for each hybrid. In addition to the genetic effects, the gene frequencies of favorable alleles would be different to magnify the expression of heterosis.

Extensive testing of populations and population hybrids through diallel mating designs has been useful to choose elite germplasm to be used as source of inbred lines. They have also helped identify potential heterotic patterns and potential populations for reciprocal recurrent selection programs. Reciprocal recurrent selection methods were proposed to genetically improve the cross of two genetically broad-based populations (Comstock et al., 1949). To improve the population cross of two maize populations in active breeding programs, the logical choice of two populations to include in reciprocal recurrent selection would be those that represent the important heterotic groups currently used to isolate inbred lines to produce hybrids (Eckhardt and Bryan, 1940). Both intra- and inter-population recurrent selection programs have been effective for changing the relative frequencies of complementary alleles that affect the expression of heterosis for grain yield (Table 10.10). The estimate of mid-parent heterosis increased with reciprocal recurrent selection (Hallauer and Carena, 2009). Average mid-parent heterosis for the five pairs of populations representing reciprocal recurrent selection increased from 7.3% for the initial population crosses to 37.4% after 6–11 cycles of selection. Mid-parent heterosis averaged across six intra-population recurrent selection programs was 40.6%. Four additional reciprocal full-sib recurrent selection programs were initiated as a consequence of these results.

Lamkey and Edwards (1991) examined the concept of mid-parent heterosis for populations and their crosses and inbred lines and their crosses. Because the classic theories of heterosis do not distinguish between heterosis expressed at the population level (designated as panmictic mid-parent heterosis) and for crosses of either inbred or partially inbred populations (designated as inbred mid-parent heterosis), they presented a theory that conceptually unifies the concept and panmictic- and inbred-heterosis. The generalized formulae they presented include level of inbreeding, frequencies of alleles, and dominance which are identical to those given by Falconer and Mackay (1996) for zero inbreeding ($F = 0$) except Lamkey and Edwards (1991) used the $F_2$ as the reference population whereas Falconer and Mackay (1996) deviated their results from the $F_2$ mean. Panmictic mid-parent heterosis was defined as the difference between the mean of the population cross and the mean of two parental populations, the panmictic mid-parent value. They also derived similar expression for panmictic-mid-parent $F_2$ heterosis and for panmictic-mid-parent $F_1$-selfed heterosis. From use of these relations, they defined the concept of baseline heterosis as the difference between the panmictic mid-parent value and the inbred mid-parent value, which is the average heterosis expected in crosses of inbred lines derived from two populations. Inbred mid-parent heterosis, therefore, is equal to baseline heterosis and panmictic heterosis. Similar expressions were developed for inbred-line–hybrid heterosis theory. As expected random mating and selfing the $F_1$ generation had the same result $[(\frac{1}{2})d]$ which leads to a
Table 10.10  Estimates of mid-parent heterosis for several recurrent selection programs conducted in temperate area maize populations that emphasized selection for increased grain yield

<table>
<thead>
<tr>
<th>Populations</th>
<th>Sources</th>
<th>Cycles of Selection</th>
<th>Response cycle$^{-1}$</th>
<th>Mid-parent heterosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>no.</td>
<td>Direct</td>
<td>Indirect</td>
</tr>
<tr>
<td>Jarvis and</td>
<td>Inter-population RS</td>
<td>Moll and Hanson (1984)</td>
<td>10</td>
<td>2.7</td>
</tr>
<tr>
<td>Indian Chief</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS10 and BS11</td>
<td></td>
<td>Eyherabide and Hallauer (1991)</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>BSSS and BSCB1</td>
<td></td>
<td>Keeratinijakal and Lamkey (1993)</td>
<td>11</td>
<td>7.0</td>
</tr>
<tr>
<td>BS21 and BS22</td>
<td></td>
<td>Menz et al. (1999)$^a$</td>
<td>6</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Menz et al. (1999)$^b$</td>
<td>6</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>Intra-population RS</td>
<td>Leaming and Midland (2001)$^c$</td>
<td>3</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Carena and CGSS (2005)$^d$</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>BS21 and BS22</td>
<td></td>
<td>Carena (2005)</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Carena (2005)</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

$^a$Populations used as testers;
$^b$Inbred lines used as testers: A632 for BS21 and H99 for BS22
$^c$Inbred progeny ($S_1 – S_2$) recurrent selection used within populations for the heterotic group suggested by Kauffmann et al. (1982)
$^d$Mid-parent-heterosis values are a result of a diallel mating design of 10 early-maturing populations across 29 environments (Carena, 2005; Carena and Wicks III, 2006). Additional data were generated by Melani and Carena (2005). Therefore, response of selection is not reported

Source: Hallauer and Carena (2009)

50% reduction in heterosis, as shown earlier by Falconer and Mackay (1996). For the inbred-line–hybrid model baseline heterosis is zero. The reason we are unable to determine proportion of heterosis between two inbred lines is either because of genetic diversity or because of inbreeding depression. Lamkey and Edwards (1991) illustrated the use of the theoretical formulae for populations and population crosses
for the studies reported by Eyherabide and Hallauer (1991) and Keeratinijakal and Lamkey (1993). For the Keeratinijakal and Lamkey (1993) report, panmictic heterosis increased over cycles of selection. Inbred and panmictic heterosis values were similar leading to a negative estimate of baseline heterosis. Somewhat different conclusions were found for the report by Eyherabide and Hallauer (1991). If only panmictic heterosis was determined, one would conclude panmictic heterosis increased because of an increase in genetic divergence of the two populations. But they determined baseline heterosis was relatively consistent across cycles of selection, whereas inbred heterosis showed a consistent increase across cycles of selection. They concluded that better designed experiments were needed to the adequacy and validity of quantitative genetic models to explain the genetic basis of heterosis.

A 4-year effort and investment to sequence the B73 maize genome was completed in 2009. However, an explicit genetic basis of the expression of heterosis for each hybrid (maize cultivars) is probably not realistic (Hallauer and Carena, 2009) even though generation of knowledge through discovery always is beneficial. Heterotic effects are unique for each hybrid and sequencing efforts on only B73 may limit the identification of useful alleles for drought tolerance and other complex traits (Carena et al., 2009). Hundreds of thousands of genotypes represent the maize genome, not just B73 or a few elite genotypes.

Due to intellectual property rights it is currently challenging to assess the level of genetic diversity available on farms. There is, however, a perceived interest to increase the genetic diversity of US maize hybrids at the farm level. One way would be to identify and exploit alternative heterotic patterns. Heterotic groups are groups of related genotypes (e.g., elite males) expressing high level of heterosis when crossed with other groups of related genotypes (e.g., elite females). As seen in Table 10.3 alternative heterotic patterns have been proposed in the 1970s and the 1980s but needed improvement before utilization. Data suggested that hybrids of lines from different germplasm sources had greater yields than hybrids of lines from similar sources. Because these studies were restricted to inbred lines from few germplasm sources only the following heterotic patterns were developed:

Reid Yellow Dent (RYD) × Lancaster Surecrop (LSC)
Iowa Stiff Stalk Synthetic (BSSS) × Lancaster Surecrop (LSC)

Tables 10.4 and 10.10 show alternative heterotic patterns have been identified in both the central and northern US and heterotic groups have been identified to enhance expression of heterosis. Alternative heterotic patterns with exceptional high-parent heterosis need to be continuously identified and exploited for different regions. How many heterotic groups are currently available? Five companies representing most of the market share were contacted in 2008 in order to better assess how limited is maize genetic diversity on farms based on heterotic patterns being utilized in full-season and short-season maize production areas. Only one out of the five major companies reported recent initiatives to bring non-adapted alleles
Fig. 10.1  Current heterotic groups being used by industry

to elite lines. Also, in one out of five cases, only one heterotic group was being used. Moreover, in 80% of the cases the same relatives were used on one side of the hybrid. This should be of concern. Short-season maize seems to be favored by use of Iodents and flints, although flints were used in only one case (Fig. 10.1).

References


Enfield, E. 1866. *Indian Corn; its Value, Culture, and Uses.* D. Appleton, New York, NY.


Wright, S. 1922. The effects of inbreeding and cross breeding on guinea pigs. USDA Bull. 1121:1–60.

Future maize genetic gains are dependent on the deployment of useful genetic diversity carried out in the public sector (Smith, 2007). In order for these gains to be significant and make impact the incorporation of unique and useful genetic diversity to breeding programs actively improving germplasm and developing cultivars is needed (Carena et al., 2009b). The most successful maize germplasm was Iowa Stiff Stalk Synthetic or BSSS (Sprague, 1946), a genetically broad-based population. Its successful spin-off line, B73 (Russell, 1972), was derived after five cycles of half-sib recurrent selection and several years of pedigree selection method of inbred line development with hybrid testing (Carena et al., 2009a). It generated billions of dollars and the benefits were shared before intellectual property was present. Even though the odds of developing successful public lines from genetically broad-based improved populations seem to be low, it only requires one to make significant impact (Hallauer and Carena, 2009).

Adequate choice of germplasm and pre-breeding is essential to increase the useful genetic diversity of maize hybrids and develop new and unique cultivars (Carena, 2008a). Without choosing the right germplasm neither traditional breeding nor modern breeding will be successful (Carena, 2008b). The use and improvement of diverse genetic resources from several sources is not often a priority within private breeding programs. Duvick (1981) defined ‘genetic diversity in reserve’ (breeding reserves) as the vast range of genetic diversity in use or nearly ready to use that is present in inbred lines, experimental hybrids, and breeding populations. The genetic diversity currently available within and among breeding programs and long-term storage can potentially provide an insurance to reduce genetic vulnerability. Public maize breeding programs provide genetic diversity in reserve (Duvick, 1981), breeding creativity, and an insurance against genetic vulnerability through continuous genetic improvement of elite genetically broad-based breeding populations (Carena and Wicks III, 2006). The genetic diversity of materials needs to be sustained to minimize the vulnerability inherent to growing uniform and closely related materials over wide areas.

In the past, the large number of active breeding programs, the introgression and incorporation of new germplasm into their programs, and the diversity of breeding methods employed reduced concerns about lack of maize genetic diversity
The number of active public maize breeding programs has significantly decreased (Frey, 1996). Public maize breeding programs are either discontinued or eroding because of the lack of federal and state funding. This loss of public support affects breeding continuity, the utilization and improvement of the maize germplasm available, and, equally important, the training of future maize breeders. Although the number of maize germplasm improvement programs in the public sector has decreased in the past 10 years, the efforts of the few remaining have increased. The future of extensive maize germplasm enhancement on a long-term basis is still uncertain though. Even though the problem has been recognized the number of public maize breeding programs enhancing germplasm and developing cultivars continues to decline. Breeding objectives, industry demands, and change in consumer preferences determine new breeding routes to explore. Without any doubt, new and useful elite germplasm and long-term genetic enhancement will enhance plant breeding programs. Most of the desirable genes in maize have not been utilized yet.

Choice of germplasm, either fortuitous or planned, plays an important role in any breeding program, whether an applied breeding program for inbred line development and/or population improvement or a selection study comparing breeding methods. There are certainly differences among breeding populations, and the particular choice of germplasm may decide the ultimate success or failure of selection. Maize breeders have always had a wealth of germplasm available for their use. Even within regions having similar environmental aspects, experience and testing soon identified certain populations (or varieties) that were better than others for use as varieties themselves, use in variety crosses, and use as source populations for the development of inbred lines as parent stocks for hybrids. Experience in the USA has shown that some breeding populations (e.g., Reid, Lancaster, Minnesota 13, and Stiff Stalk Synthetic) yield a greater frequency of usable inbred lines than others (e.g., Hickory King, Krug, and Corn Borer Synthetic 1). Differences among breeding populations occur because of the original genes and past selection that created an assemblage of genes in the greater frequencies that are desired in modern hybrids. Selection may have been as intense and effective for other populations, but perhaps the original germplasm in the populations did not include the genes necessary to meet the standards of the breeder. If the genes are not present, the efforts of the breeder will be futile regardless of the patience and experimental techniques used. Hence, maize breeders have two important but separate decisions to make in developing their breeding programs:

2. Choice of breeding procedure.

Maize is an extremely diverse genus, having many morphological and biological differences. Maize is a monoecious plant that has separate male and female inflorescences and because it is essentially 100% cross-pollinated through wind movement the opportunities for intercrossing are ample. Maize has been easily adaptable (e.g., maize flowering has been a very simple trait for breeders to improve at a rate of 2–3
days per year, see previous chapters) and, therefore, it is grown from 58°N latitude without interruption through the temperate, subtropical, and tropical regions of the world to 40°S latitude. Growing of maize in the Northern Hemisphere (e.g., Canada, northern Europe/USA, and Russia) to the Southern Hemisphere (e.g., Australia, South Africa, and Argentina) permits selection of maize types for each ecological niche. In the Andean region, maize is grown from sea level to elevations over 3,808 m (above Lake Titicaca in Peru) and from areas with less than 25.4 cm of rainfall (such as the Guajira Peninsula of Colombia) to over 1,016 cm (Department of Choco on the Pacific Coast of Colombia) (Grant et al., 1963). Maize is grown in all states of the USA and in every other important agricultural area of the world. The extent of maize culture, probably greater than any other cultivated crop, is very important in the world economy for food, feed, fiber, and lately for fuel.

Maize breeders have been increasingly cognizant of the importance of genetic diversity of germplasm in the 20th century. The expression of heterosis, whatever its genetic basis, depends on the differences in allele frequencies of the parental stocks, whether varieties or inbred lines are used to produce the crosses. Because of the heterotic responses observed in crosses of maize, breeders have emphasized crossing of parental stocks that were derived from different breeding populations. Initially, the concern of diversity of source populations, which were usually open-pollinated varieties, was either for crossing to produce variety crosses or to initiate inbreeding for developing inbred lines. Past selection, either natural or human, in different regions developed germplasm that had distinctive phenotypic features (e.g., Wallace and Bressman, 1925) and different gene frequencies for different traits as evidenced by the expression of heterosis in crosses (see Chapter 10).

The concern of genetic diversity of germplasm has received emphasis because of

1. the rapid shift from double crosses to simpler types of hybrids,
2. the Bipolaris maydis outbreak on hybrids produced on T-cytoplasm in the USA in 1970, and
3. the perceived reduced genetic diversity in commercial hybrids, especially in the 21st century with the offer of hybrids differing only in transgenic events (e.g., GEM project).

It was shown in Chapter 9 that the use of simpler types of hybrids does not seem to have any serious disadvantages compared with double crosses. A series of surveys sponsored by the American Seed Trade Association (Sprague, 1971; Corn, 1972; Zuber, 1975; Darrah and Zuber, 1986) and reports by Mikel (2006, 2008) and Mikel and Dudley (2006) have indicated extensive usage of a few publicly developed inbred lines in hybrids. The first two surveys related primarily to the use of inbred lines in double crosses while the last three surveys were taken after the rapid shift from double crosses to single crosses. Double crosses, because they are produced from the union of gene arrays of two single crosses, were more variable and provided a level of genetic diversity within each field and among fields that would be expected to be greater than that resulting from the use of single crosses. It seems, therefore, that genetic diversity of hybrids produced and grown
was reduced substantially in the US Corn Belt after the transition to single crosses. Level of genetic uniformity, however, would not approach the extensive use of only one cytoplasm in producing hybrid seed. Because maize is an annual crop the problem of uniformity seems to be less serious than for a perennial crop, provided that a source of germplasm and of new breeding materials is available to shift, in a short period of time, the hybrids of varieties cultivated for grain production. The 1970 \textit{B. maydis} debacle did, however, emphasize the potential seriousness of the problem of reduced genetic diversity of our germplasm. Subsequent reports (Sprague, 1971; Genetic vulnerability of major crops, 1972; Recommended actions and policies, 1973; Lonnquist, 1974; Brown, 1975) have discussed the potential seriousness of reduced genetic diversity, outlined steps that are necessary to correct the situation, and recommended possible avenues of research to reduce genetic uniformity of breeding materials. Although breeders have at their disposal a nearly unlimited diversity of maize germplasm, the main problem arises from usage of parental inbred lines in hybrids. If a single cross of two inbred lines is superior to other single crosses (e.g., B73 × Mo17), economic considerations and competition among seed producers force widespread usage of either one or a few specific hybrids. Troyer (2004), Mikel and Dudley (2006), and Mikel (2006, 2008) have shown, however, that a few elite genotypes are persistent and continue to have important roles in breeding programs because of the recycling methods used to make consistent, incremental genetic improvements with each cycle of selection of elite line crosses.

11.1 Origin of Maize

Maize is a member of the grass family placed in the tribe Maydeae (also called Tripsaceae by some authorities, e.g., Hitchcock, 1935). The tribe Maydeae includes seven genera, two that are native to the Western Hemisphere and five that are native to Asia. Each genus has the common trait of separate male and female inflorescences on the same plant. The two genera native to the Western Hemisphere include Zea and Tripssacum; in some instances a third genus, \textit{Euchlaena}, is considered. Currently, the annual weed teosinte previously classified as \textit{Euchlaena mexicana} is included in the same genus as maize (\textit{Zea mexicana}). A perennial tetraploid (2n = 40) form of teosinte (\textit{Z. perennis}), thought to be derived from \textit{Z. mexicana}, is often classified as a separate species, and \textit{Z. perennis} is usually not considered in the origin of any of the other living species. The rediscovery of perennial diploid teosinte (\textit{Z. diploperennis} [Gramineae]) in southern Jalisco, Mexico, may provide valuable clues to the evolution of \textit{Zea} and to the origin of \textit{Z. perennis} (Iltis, 1979).

Annual teosinte plants resemble maize plants but commonly are more slender and have several stalks (or tillers) per plant. Teosinte ears are smaller than maize ears, usually having only five or six seeds per row and two rows per ear. Each kernel is enclosed by a horneous shell borne on a hardened, brittle rachis. Annual teosinte and maize each have 10 pairs of chromosomes. The chromosomes of teosinte resemble those of maize but tend to have more knob-like structures than those of maize. Maize and teosinte cross relatively easily and produce fertile offspring.
Tripsacum includes several perennial species and the resemblance between maize and Tripsacum species is much less than between maize and teosinte. The relationship between Tripsacum and maize seems to have been discovered about the same time as the relationship between maize and teosinte. Male flowers are located on the upper portions of the tassel and the female flowers on the lower portion. Each seed is enclosed by a horny covering, but seeds are not covered by leaves or husks as for maize and teosinte. Tripsacum species have chromosome numbers that are multiples of 36 and, therefore, special techniques are usually required for crossing Tripsacum and maize. All maize–Tripsacum crosses so far have been male sterile and the male sterility persists after several generations of backcrossing to maize.

The five Asiatic genera of Maydeae have not been studied as extensively as the two genera native to the Western Hemisphere. Coix is the only Asiatic genus ever seriously considered as a possible ancestor of maize. Relatively little is known about the other four Asiatic genera, Chionachne, Polytoca, Sclerachne, and Trilobachne. Chionachne and Sclerachne have 20 chromosomes, Polytoca has 40, different species of Coix have 10, 20, and 40 chromosomes, and the chromosome number is unknown for Trilobachne. Coix is the only Asiatic genus that has been successfully crossed with maize. Maize also has been crossed with sugarcane (Saccharum officinarum), but apparently sugarcane crosses with many grasses without regard to degree of relationship (Goodman, 1965a).

The origin of maize has been studied extensively, but the putative parents of cultivated maize as we know it today are still conjectural. Extensive literature indicates the extent of the study of the origin of cultivated maize, but the issue is still unsettled. Reviews by Weatherwax (1955), Goodman (1965a), Mangelsdorf (1974), Galinat (1977), and Wilkes (2004) illustrate the differences of opinion regarding the origin of maize, but there seems to be general agreement that it originated in the Western Hemisphere.

Historically, four hypotheses regarding the possible origin of maize were proposed:

1. Maize, teosinte, Tripsacum, and perhaps some of the Andropogoneae descended from a common, extinct ancestor native to the highlands of Mexico or Guatemala (Weatherwax, 1955).
2. Maize originated from a cross between two species, perhaps Coix and Sorghum, each with 10 chromosomes (Anderson, 1945).
3. The tripartite hypothesis of Mangelsdorf and Reeves (1939) postulated that (a) wild maize was a form of pod corn native to the lowlands of South America, (b) teosinte originated from crossing cultivated maize and Tripsacum in Central America, and (c) modern varieties of maize arose from crosses between maize and Tripsacum or teosinte.
4. Maize was derived from teosinte by direct human selection (Beadle, 1939).

These four divergent theories of the origin of maize have stimulated research in an attempt to resolve the issue. The hypotheses have been modified in some instances
as additional evidence became available (Goodman, 1965a; Galinat, 1977; Wilkes, 2004).

Weatherwax’s proposal is the simplest hypothesis and has received support (e.g., Brieger et al., 1958) for that reason. He does not consider the oriental Maydeae to be closely related to maize. Weatherwax (1955) described the traits that would be necessary in a wild form of maize and concluded that teosinte was not wild maize. He suggested that wild maize probably became extinct soon after the Native Americans began growing maize; that the chromosome number in *Tripsacum* probably prevented crossing between *Tripsacum* and maize or teosinte; and that maize and teosinte, as well as *Tripsacum*, arose in mutually exclusive isolation because of the ease of crossing between the two species. Weatherwax’s proposal has been criticized for relying on three separate areas of origin; for assuming the development of modern races of maize from primitive races of maize and wild maize; and because of the presence of the heterochromatic knobs in maize, which he attributes to contamination from teosinte.

Anderson (1945) hypothesized that the origin of maize was in southwestern Asia; hence the suggestion that perhaps maize originated from a cross of *Coix* and *Sorghum*. The hypothesis (*Coix* × *Sorghum* = maize), however, has not been considered seriously. Objections to this hypothesis have arisen as to when and how the genera of Maydeae became distributed in southern Asia and Central America because of the discovery of ancient maize pollen and the little genetic evidence that maize originated by doubling a basic chromosome number of 10. Recent crosses of maize and *Coix* may provide additional information on the validity of Anderson’s hypothesis.

The tripartite theory of Mangelsdorf and Reeves (1939) has stimulated research that either supported or detracted from the validity of the postulated theory. It has probably received more attention than the others. The first aspect of the tripartite theory was modified. The original form of pod corn as described by Weatherwax was changed to a pod corn type in which each kernel is only partially enclosed by a small husk. Also, the center of origin was changed from the lowlands of South America to Mexico. Teosinte was considered to have arisen from crosses of *Zea* and *Tripsacum* because teosinte seemed intermediate for several traits. This second aspect has been questioned because

1. maize and *Tripsacum* are so variable that teosinte seems more uniform by comparison,
2. some claim that teosinte is not intermediate to maize and *Tripsacum*, and
3. hybrids of maize and *Tripsacum* are sterile, forming a barrier for the exchange of genetic factors from *Tripsacum* to maize.

The third aspect of the tripartite theory has received greater acceptance than the first two, but there has been some disagreement on the relative influence of *Tripsacum* and teosinte on the development of maize. Because teosinte and maize are nearly identical, genetically and cytologically, the general consensus is that teosinte has contributed more to the development of maize than has *Tripsacum*. The validity
of the tripartite hypothesis for explaining the origin of maize has been abandoned by its principal proponent. Mangelsdorf (1974), after spending nearly 30 years investigating the origin of maize, concluded that the hypothesis was not adequate. Electron microscope studies of the pollen of maize, teosinte, Tripsacum, and maize–Tripsacum hybrids by Mangelsdorf’s colleagues showed convincingly that teosinte was not a hybrid of maize and Tripsacum (Mangelsdorf, 1974).

The hypothesis that maize was derived directly by human selection from teosinte was suggested by Beadle (1939). Subsequent studies by Galinat (1970, 1971, 1975), Iltis (1970, 1972), de Wet and Harlan (1971, 1972, 1976), Beadle (1972, 1977), and Kato (1975) have presented additional evidence. Beadle (1977) studied F2 and backcross generations of crosses between teosinte and primitive maize types (Argentine pop, Chapalote, and Chalco teosinte). From genetic analysis of 16,000 segregants, he found the occurrence of parental types in frequencies of about 1 in 500, suggesting a relatively few independently segregating genes. Demonstration trials also showed that teosinte yields were comparable to those of wild wheat in the Near East. Sufficient quantities of seed could be harvested to be used for human food. Kahn (1985) provided an interesting account of the debate among the different scientists for the relevancy and possible impact of teosinte on the origin of maize.

Galinat (1977) and Wilkes (2004) gave a detailed account that reviewed the evidence currently available for the origin of maize. Comparisons of maize with its relatives (or putative parents) were made on the basis of genetics, cytology, floral structures, fossil evidence, and the relative contributions of maize relatives to maize improvement. Galinat summarized as follows: ‘Of the various hypotheses on the origin of maize, essentially only two alternatives now remain as viable options:

(la) Present-day teosinte is the wild ancestor of maize,
(lb) A primitive teosinte is the common wild ancestor of both maize and Mexican teosinte,
(2) An extinct form of pod corn was the ancestor of maize with teosinte being a mutant form of this pod corn.’

The four hypotheses for the origin of maize were not always incompatible in their features, and as additional evidence became available it is only natural that modifications and revisions were made in the original hypotheses. The issue is not settled, but it seems that teosinte was important in the evolution of maize and offers greater opportunities than Tripsacum as a source of genetic variation for maize breeding. Tripsacum has possibilities but the technical problems of introducing its germplasm into maize germplasm are much greater. For future studies, Goodman (1965a) feels that greater emphasis should be given to the Asiatic genera of Maydeae and Galinat (1977) feels that the greater genetic variability of the nine or so species of Tripsacum offers potential usefulness for maize improvement. Rediscovery of Z. diploperennis will provide geneticists and breeders potentially valuable germplasm for gene transfer (Iltis, 1979). Future research with Z. diploperennis may have important implications in studying the origin of maize, but practical uses in maize improvement still remain in the future.
As of today, geneticists accept that maize is solely derived from teosinte (Z. mays spp. parviglumis), specifically from a population of Balsas Teosinte (Wilkes, 1967; Iltis, 2006).

11.2 Classification of Maize Germplasm

If one considers only the obvious phenotypic differences, the range of variability that exists among races, varieties, hybrids, and inbred lines of maize for plant type, ear type, tassel type, and maturity is impressive. The obvious phenotypic differences are relative to the germplasm often included annually in breeding nurseries, which may vary from adapted germplasm to introduced germplasm, but significant differences are always present. Because of the range in latitude and altitude in which maize can be grown, it is little wonder that so many different types of maize have been developed. It seems that maize has been cultivated for 5,000 years; consequently its use in satisfying food, fuel, feed, and fiber as well as cultural needs of Native American settlements created a vast array of germplasm. Development of the cultures of the different groups of peoples, their migrations, discovery of the Western Hemisphere, and the subsequent movement of Europeans into it also were important factors in creation of a diversity of maize germplasm. Because of the cross-pollination accompanied by a continuous interchange of genes among populations, additional pools of genetic variation have been created by the movement of peoples. Subsequent selection, both natural and artificial, developed germplasm that was often quite different in phenotype and genotype from the original parental germplasm.

The vast array of maize germplasm was obvious to students, fanciers, taxonomists, botanists, and breeders of maize, but no natural classification was attempted until the 1940s. Sturtevant (1899), one of the first to make the attempt, classified the maize germplasm known to him into six main groups, five of which were based on endosperm composition. This system was generally used without modification for over 40 years; little other interest and activity in classification were shown.

N. I. Vavilov and his associates collected a large number of specimens from different areas of the world and they concluded the center of origin of maize was in Central America because of the variability of maize types in that area. As a follow-up of those collections, Kuleshov (1933) classified maize by endosperm types in the following groups:

1. Z. mays indurata – flint
2. Z. mays amylacea – floury
3. Z. mays indentata – dent
4. Z. mays everta – popcorn
5. Z. mays saccharata – sweet
6. Z. mays amylea saccharata – starchy-sugary
The classification of Kuleshov was similar to Sturtevant’s because in some instances only a one gene difference was needed to change the classification. The classification was satisfactory for kernel type, but it was not indicative of the germplasm’s morphological and polygenic differences for other traits. Geographical distribution of the eight systematic groups listed by Kuleshov follows:

1. Flint maize was distributed throughout the Western Hemisphere, but its greatest importance seemed to be in northern and southern frontiers of maize growing.
2. Flour-type maize occurred south of the northern range of flints in North America and in the southwestern states of the USA, and it predominated in the Andean valley of southern Colombia, Peru, and Bolivia; the greatest diversity occurred in Peru.
3. Dent varieties were predominant in what is called the US Corn Belt and in some areas of Mexico. The dent group did not seem to occur in the aboriginal culture of South America, but the greatest diversity seemed to occur in the central and southern states of Mexico.
4. The popcorn types were collected in several countries and localities but, except for commercial production, had not moved into North America.
5. Sweet maize was collected principally in the central and northeastern regions of the USA and was nearly absent in the south and the tropics.
6. Nearly all the starchy-sugary types were collected in Bolivia and Peru with the greatest diversity found in Peru.
7. The waxy types seemed to be restricted to eastern Asia.
8. No fixed geographical area was identified for the tunicata types. The Tu allele is never found in any particular type, and the spontaneous occurrence of tunicata types can occur in different areas.

The method of classifying maize germplasm by endosperm type was not satisfactory, as noted by Sturtevant (1899). But it was not until Anderson and Cutler (1942) investigated the range of variability among collections of maize germplasm and developed the concept of ‘races of maize’ that vigorous effort was given to classifying maize germplasm. The definition of a race by Anderson and Cutler was ‘For the classification of Z. mays we shall define the word race as loosely as possible, and say that a race is a group of related individuals with enough characteristics in common to permit their recognition as a group.’ They continued by stating, as Hooton (1926) has said in his discussion of racial analysis, “Races are great groups and any analysis of racial elements must be primarily an analysis of groups, not of separate individuals. One must conceive of race not as the combination of features which gives each person his individual appearance, but rather as a vague physical background, usually more or less obscured or overlaid by individual variations in single subjects and realized best in a composite picture.” Genetic and phenotypic differences were considered in their definition of a race. Anderson and Cutler (1942) thus
stated that ‘a race or sub-race is defined as a number of varieties with enough characters in common to permit their recognition as a group; in genetic terms it is a group with a significant number of genes in common.’ Thus a more natural system rather than an artificial description was developed. Anderson and his colleagues used this descriptive definition for classifying maize germplasm and the concept has been used extensively. If used only for cataloging, inventorying, or storing, the method of classifying germplasm suggested by Sturtevant is adequate, but the classification suggested by Anderson and Cutler is certainly more useful in attempting to derive and trace the origins of different races.

The concept of race is not easily understood. Wellhausen et al. (1952) and Brieger et al. (1958) discussed the concept of a race and how races of maize may have originated. It is generally agreed that races exist and are characterized by complexes of traits that make one race distinguishable from another. Levels of differentiation among races are not always the same, but races seem to maintain themselves for many generations without losing their identity. Brieger et al. (1958) elaborated on the definition of a race as follows: ‘We may define as a race any group of populations having sufficient number of distinctive characters in common, maintaining itself through panmictic reproduction within populations, and occupying definite areas.’ This definition was not in conflict with the one given by Anderson and Cutler (1942), but the concise definition does describe a race as a random mating population possessing definite phenotypic and genetic characteristics. The difficulty of precisely defining a race in general terms was equated with the problems of a precise definition for a species as there usually are exceptions to the general definition.

Races of maize have arisen, but it is not always clear what mechanisms were involved and how the races maintained their integrity through many generations of reproduction. Distinctive races have evolved in different regions following the distribution of primitive maize several thousands of years ago. Initially, frequent mutations and isolation mechanisms (geographical, flowering, and gametophyte factors) must have played a prominent role. Superimposed on the evolutionary trends were the activities of peoples by migration and thus the movement of germplasm to different geographical areas. Native Americans contributed to the maintenance of different races by isolating populations with special characteristics, particularly ear and grain traits (e.g., grain color) for use in their ceremonies and/or other purposes. Races of maize arose simply because of artificial and natural selection pressures that caused changes in allele frequencies in successive generations of propagation. Wellhausen et al. (1957) cited evidence that maize under domestication is potentially a self-improving species, as evidenced by the increasing size of the ear for the past 4,000 years. The crossing of distinct races, probably unintentional initially, enhanced productivity, and new races evolved because of hybridization. Wellhausen and his colleagues concluded that mutation and racial hybridization were the two important evolutionary factors for the development of races in Mexico and Central America.

Brieger et al. (1958) also emphasized that isolating mechanisms must exist for races to have maintained their racial characteristics for many generations of inter-mating, which, in some examples cited, were not obvious. Races of maize were
considered to have arisen by selection of mutant genes accompanied by favorable modifier complexes and by a similar procedure after hybridization of previously existing races. Both ways of explaining the development of races were similar to those of Wellhausen et al. (1952), but Brieger et al. emphasized that the intermediacy of presumed synthetic races created by hybridization may not be valid for postulating the putative parent races. Synthetic populations formed by crossing two parents tend to be intermediate to the two parents, but Brieger et al. question selecting putative parents that possess characters differing equally in opposite directions of the assumed synthetic race as one should not expect that the synthetic race will be intermediate in all or even most of the characters.

The concept of race was a successful attempt to classify maize germplasm. Races were characterized by differences in quantitative characters that are often variable within races (Wellhausen et al., 1952). Use of quantitative traits rather than simply inherited traits was a more natural system of classification, but expressions of quantitative traits are subject to environmental biases and breeding information was usually lacking to assist in the classification. Because of the extent of germplasm available for study, other methods of classification are being investigated to assist in classifying races. Goodman and Paterniani (1969) listed three ways that environmental biases may be reduced:

(a) Evaluate the germplasm in several environments and use average values of traits over environments.
(b) Evaluate the germplasm in several environments and determine similarities of responses within each of the environments.
(c) Limit comparisons to those characters that have the least environmental bias relative to size of differences among means.

Goodman (1967, 1968), Goodman and Paterniani (1969), Bird and Goodman (1977), and Goodman and Bird (1977) have used numerical taxonomy techniques in an attempt to identify characters that fulfill condition (c) and to identify how the techniques relate to previous race classifications. Characters least affected by environmental factors and their interactions with environments were reproductive characters (e.g., ear and kernel characters) that had components of variance greater than the sum of corresponding components of variance for years and race by year interactions (Goodman and Paterniani, 1969). Vegetative characters tended to have greater interactions with environments and tassel characters were intermediate to the reproductive and vegetative characters. Hence reproductive characters seemed to be better indicators of racial differences than vegetative characters.

Preliminary information on numerical methods of classification shows that multivariate methods (principal component analysis, factor analysis, canonical variate analysis, cluster analysis using unweighted variables, etc.) can provide additional information in sorting out relations of races of maize. Use of the numerical methods of taxonomy, however, is very complex and only limited efforts have been directed to their use. The feasibility of different methods of classification, as well as more conventional classification procedures, depends on the choice of characters.
measured and how they are influenced by environmental factors. There has been a consistent relation in classification of races between multivariate analyses and conventional procedures. As classification procedures and techniques become developed, as additional data are collected for different morphological and physiological traits, and as additional breeding information becomes available, classification of races should become more refined (Paterniani and Goodman, 1977). Adjustments will be made for duplications and classifications will become more subjective. Collectively the information will be valuable to breeders in the range of variability among maize races and will help them to recognize patterns of variation in choice of germplasm and in creating new groups of genetic diversity.

11.3 Races of Maize in the Western Hemisphere

The first extensive race description of a comprehensive collection of maize germplasm was reported by Wellhausen et al. (1952) in Mexico. Similar types of programs were later initiated throughout the Western Hemisphere under the auspices of the Committee on the Preservation of Indigenous Strains of Maize within the National Academy of Sciences National Research Council (NAS-NRC). Because new hybrids and varieties had rapidly replaced open-pollinated varieties in the USA, the NAS-NRC was concerned with collection and preservation of maize varieties indigenous to the Western Hemisphere. Much of the original germplasm in the USA in the form of open-pollinated varieties was lost even though over 800 unique open-pollinated varieties were present (Sturtevant, 1899). Maize is one of the basic food plants in the Western Hemisphere; and its great diversity, resulting from thousands of years of domestication, was considered one of the important natural resources of the Western Hemisphere. It was felt that rapid development of communication, travel, and breeding programs would result in the same fate for indigenous varieties and races of other countries. Hence, an effort to collect, study, and preserve the vast reservoir of genetic variability was initiated.

The format of collecting, studying, and classifying used by Wellhausen et al. (1952) was followed in nearly all instances. Characters used in classifying the maize collections comprised four principal categories:

1. Vegetative characters of the plant response to altitude, height, total leaf number, number of leaves above the ear, width of ear-bearing leaf, venation index, and internode patterns.
2. Tassel characters: Tassel length, peduncle length, length of branching space, percentage of branching space, percentage of secondary and tertiary branches, and total number of branches.
3. External and internal Ear characters: Ear diameter, length, and row number; shank diameter and length; number of husks; kernel width, thickness, length, and denting; cob diameter; rachis diameter; cob/rachis index; rachilla length;
11.3 Races of Maize in the Western Hemisphere

rachilla/kernel index; glume/kernel index; cupule hairs; rachis flag; lower and upper glume traits; rachis induration; and teosinte introgression.

4. Physiological, genetic, and cytological characters: Number of days from planting to anthesis, pubescence of leaf sheath, plant color, mid-cob color, chromosome knobs, and B-chromosomes.

The characters measured described the entire genetic constitution of the germplasm under consideration rather than only Sturtevant’s endosperm differences.

Studies conducted on the races of maize in the Western Hemisphere are summarized in Table 11.1.

Maize races are listed in chronological order of publication and show the number of races described in each country and the number of collections studied in arriving at the racial classifications. In addition, each report gives a description of ecological conditions, maize cultures, and methods of collection and classification. A wealth of information is included in the reports, and they probably represent the only detailed repository of information relative to the races of maize. A total of 285 races was described. There was some overlapping of races for different countries and regions, but classification of the collections differentiated distinctive races. It was fortunate that the studies were undertaken when they were (although sooner would have been preferable) because considerable hybridization of races was occurring by interchange of germplasm. The classification identified where and when hybridization had occurred and the putative parents involved. Collection and classification were effective in developing the lineage of many races.

Brown and Goodman (1977) and Goodman and Brown (1988) summarized the reports listed in Table 11.1. They provided what seemed to be well-defined racial groupings. The groupings provide information on relationships among races and should be helpful to breeders in selecting germplasm. A re-examination of races described by authors listed in Table 11.1 indicates that many of the races were duplicates and that about 130 more or less distinct racial complexes make up the aggregate maize germplasm of the Western Hemisphere. Most of the reports listed in Table 11.1 were intended to be only preliminary because little or no breeding information (inbreeding and hybridization) was available to assist in the classification of the races.

Brown and Goodman (1977) list and describe nine races for the USA. According to reports listed in Table 11.1, variability of maize germplasm in the USA is considerably less than in other areas (e.g., Wellhausen et al., 1952; Roberts et al., 1957; Brieger et al., 1958; Grobman et al., 1961). Fewer races were identified, but much of the maize germplasm of the USA was lost before it became apparent that it would be useful in breeding programs. Development of and interest in inbred lines and hybridization caused maize breeders to virtually ignore the basic germplasm sources in the USA. Iodent, for example, was a strain of Reid Yellow Dent that was developed by L. C. Burnett at Iowa State University in the early 1900s (Wallace, 1923). M. T. Jenkins included Iodent in his 1922 breeding nursery from which I205 (Idt) and 18 other Iodent lines were developed. But no record or knowledge of seed of the Iodent variety being available after the
<table>
<thead>
<tr>
<th>Source</th>
<th>Areas</th>
<th>Number of collections</th>
<th>Number of races described</th>
<th>Source Areas</th>
<th>Number of collections</th>
<th>Number of races described</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellhausen et al. (1952)</td>
<td>Mexico</td>
<td>2,000</td>
<td>32</td>
<td>Total</td>
<td>2,000</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Subdivisions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ancient indigenous</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pre-Columbian exotic</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Prehistoric mestizos</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Modern incipient</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Poorly defined</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Hathaway (1957)</td>
<td>Cuba</td>
<td>—</td>
<td>7</td>
<td>Commercial</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Domestic</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Roberts et al. (1957)</td>
<td>Colombia</td>
<td>1,999</td>
<td>23</td>
<td>Primitive</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Probably introduced</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Colombia hybrid races</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Wellhausen et al. (1957)</td>
<td>Central America</td>
<td>1,231</td>
<td>13</td>
<td>Primitive</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Exotic and derived</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Brieger et al. (1958)</td>
<td>Brazil</td>
<td>3,000$^a$</td>
<td>52</td>
<td>Argentina</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eastern slopes of Andes</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Under the Capricorn</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Amazon Basin</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Ramirez et al. (1960)</td>
<td>Bolivia</td>
<td>844</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown (1960)</td>
<td>West Indies</td>
<td>135</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timothy et al. (1961)</td>
<td>Chile</td>
<td>39–114</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grobman et al. (1961)</td>
<td>Peru</td>
<td>1,600</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timothy et al. (1963)</td>
<td>Ecuador</td>
<td>675</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grant et al. (1963)</td>
<td>Venezuela</td>
<td>685</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown and Goodman (1977)</td>
<td>USA</td>
<td>—</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brandolini (1969)</td>
<td>Europe</td>
<td>6,000</td>
<td>11$^b$(33)</td>
<td></td>
<td>6,000</td>
<td>285</td>
</tr>
</tbody>
</table>

$^a$Obtained from Paterniani and Goodman (1977), who also describe 91 populations belonging to 19 races and 15 sub-races.

$^b$Eleven races described by Leng et al. (1962) and 33 groups listed by Pavlicic (1971) not included in total of 285.
mid-1930s is known. The only known source of Iodent germplasm is BS30, a synthetic variety constructed by intermating 19 Iodent lines that Jenkins developed from his sampling of Burnett’s Iodent variety in his 1922 breeding nursery (Hallauer, 1995). Similar fate occurred for other open-pollinated varieties and is unfortunate that extensive collections were not classified and preserved in the early 1900s. It is also unfortunate that unique inbred line derivatives have been lost due to either closing of public breeding programs or reducing them to programs lacking an applied component for germplasm improvement and cultivar development.

Paterniani and Goodman (1977) summarized relative proportions of described races based on adaptation to elevation and endosperm type. About 50% of the races was adapted to low altitudes (0–1,000 m), about 10% was grown at intermediate altitudes (1,000–2,000 m), and about 40% was adapted to higher elevations (greater than 2,000 m). Based on endosperm type, the following classifications of the races were made: About 40% floury, about 30% flint, slightly more than 20% dent, about 10% popcorn, and about 3% sweet corns. No classification was given relative to endosperm types adapted to low, intermediate, and high altitudes. Adaptation to elevation was considered mainly a result of natural selection, whereas distribution of endosperm types was primarily due to human preferences.

11.4 European Races of Maize

Maize was introduced in Europe by Christopher Columbus and was first cultivated in fields near Seville, Spain, during 1492 (Brandolini, 1971). During the next four centuries, movement of maize germplasm from the Western Hemisphere to Europe continued intermittently and at different rates. Continued influx of germplasm from the Western Hemisphere resulted in a range of variation available for use, but introduction of improved varieties and hybrids from the US Corn Belt rapidly led to the disappearance of many of the adapted varieties either by substitution or intercrossing with previously introduced germplasm. Consequently, a repetition of the US experience occurred in Europe. Many of the varieties and races selected for environmental conditions of Europe were lost because of interest in recently improved introductions from the US Corn Belt.

Brandolini (1969, 1971) and Leng et al. (1962) have described briefly the distribution of maize in Europe and surrounding regions after it was first grown in 1492. Brown (1960) suggested that the two West Indies races, Coastal Tropical Flint and Early Caribbean, may have been the original maize germplasm introduced into Europe after the voyages of Columbus but Leng et al. (1962) reported that maize types currently available in southeastern Europe bear little resemblance to these two races. Maize was poorly adapted to environments of Spain, but because of repeated collections by the explorers of the Western Hemisphere, germplasm was continuously introduced to the European continent. Four hundred years of selection developed varieties that were adaptable to the broad spectrum of environmental conditions, from the arid conditions surrounding the Mediterranean Sea to the short
growing seasons of northern Europe. Introductions of different germplasms had different impacts on the composition of the maize gene pools, and natural and artificial selection developed varieties with specific fitness to the new environments. Because of the necessity of surviving sea transport, it seems that flint and popcorn varieties played a prominent role in the early European germplasm, especially in southern Europe. Eventually the Northern Flint and Southern Dent races were introduced in the central part of Europe as a result of English and French explorations in North America. Later, about 1900, the Corn Belt Dent race became an important part of the European germplasm. The different stages of introduction, hybridization of introduced germplasm with previously introduced germplasm, and selection of types to meet the wide range of environmental conditions in Europe created a complex array of European germplasm.

Preliminary classifications of European germplasm have been given by Leng et al. (1962) for southeastern Europe and by Brandolini (1969) for the European continent. Leng et al. (1962) concluded from their studies that at least 11 maize races occur in southeastern Europe:

1. Small-eared Montenegrin flints: Probably a direct derivation of the introductions from the Western Hemisphere and with ear type similar to that of the Andean race Amarillo de Ocho.
2. Small-kerneled flints: Closely resembling South American pearl popcorns in plant and ear characteristics.
3. Eight-rowed (Northern) flints: Typical of the Northern Flints described by Brown and Anderson (1947) and probably direct introductions from the USA.
4. Mediterranean flints: Relatively rare and seemingly unlike any of the races described in the Western Hemisphere.
5. Derived flints: Seemingly resulted from hybridization of different races, e.g., hybridization of the Mediterranean flints with other flint varieties.
7. Large-kerneled dents: May have resulted from direct introductions or evolved from hybridization between dent varieties and eight-rowed flints.
9. Corn Belt Dents: Direct importations from the USA in 1890–1910, constituting the major source of germplasm in southeastern European breeding programs.
10. Derivatives of hybrids between flint and dent races due to hybridization.

Leng et al. (1962) emphasized that European germplasm may be quite useful to breeders in temperate zone regions because it is adapted to temperate zone climatic conditions and day lengths. The most common problem encountered by breeders in the US Corn Belt and other temperate zone regions from the introduction of Central and South American germplasm is its adaptability and photoperiod response. The
introduced germplasm usually requires several (5–10) years to adapt to Corn Belt conditions before it can be seriously worked into the breeding programs. European germplasm would be more adaptable to the US Corn Belt, but the gain from adaptability may be offset by the limited new genes introduced in the program. Development of the races in Europe has spanned a much shorter time than development of the old races in the Western Hemisphere. Use of European germplasm would probably have greater short-term gains, but total genetic gain may be greater with use of germplasm from outside the temperate zones.

Sanchez-Monge (1962), Brandolini (1969, 1971), Brandolini and Avila (1971), and Pavlicic (1971) have given preliminary reports on classification of maize germplasm for southern Europe and Mediterranean areas. Collection, classification, conservation, and exchange of germplasm for this area were conducted by the Southern Committee of the European Association for Plant Breeding Research (EUCARPIA). About 6,000 samples were collected and 3,260 samples have been studied (Brandolini, 1969). Major goals of the collections were to determine similarities or differences among representative samples collected from the different countries and to obtain knowledge about genetic mechanisms involved in adaptation to different areas. The collections were more extensive and included a greater area than those reported by Leng et al. (1962). It was found that the types of maize collected and studied have many traits in common, suggesting a common origin. For example, the small-eared Montenegrin flints described by Leng et al. were similar to the Poliota varieties of Italy. Pavlicic (1971) presented some preliminary data on the collections that showed

(1) considerable similarity among countries for the flint types,
(2) the greatest variability among varieties in Italy, with Yugoslavia second,
(3) considerable overlap of types from Italy and Yugoslavia, and
(4) some similarities in the Yugoslavian types to the types from Romania, Bulgaria, and the USSR.

From the same collection of germplasm, Brandolini (1971) found that the frequency of chromosome knobs was consistently low. Brandolini (1969) supported Leng et al. (1962) in that southern European races seem to possess germplasm that is of value to breeding programs in temperate regions. Although contamination by hybridization with recently introduced germplasm has occurred, a large reservoir of germplasm is available in Europe for breeding programs in temperate regions. Different selection pressures have been applied to different areas of Europe, so different races and varieties include genes that contribute to disease and insect resistance, drought tolerance, short growing seasons, spring cold tolerance, and low moisture at harvest. So far, however, movement of germplasm from Europe to the US Corn Belt has been minimal but public by private partnerships are extensive. Integration of elite dent and flint lines has been common for forage and grain hybrids in Europe. Genetic diversity estimates and testcross information on European lines have been reported elsewhere (Messmer et al., 1992; Schon et al., 1994; Lubberstedt et al., 1997a,b; Lubberstedt et al., 1998; and Lubberstedt et al., 2000).
11.5 US Corn Belt Germplasm

Although study of the maize germplasm of the USA was generally initiated too late, descriptions of the array of varieties available before acceleration of inbred line and hybridization breeding programs are available. Most of the descriptions of varietal germplasm were not systematic or directed to origin or lineage of the varieties. Brief descriptions of many of the varieties used to study heterosis observed in variety crosses were presented (see Chapter 10), where observed heterosis of varietal crosses would be a measure of genetic diversity among parental varieties. Wallace and Bressman (1925) and other similar types of references and experiment station bulletins (e.g., Atkinson and Wilson, 1914) provide descriptive accounts of some of the open-pollinated varieties but they are neither extensive nor complete. Similar to racial descriptions of reports listed in Table 11.1, there were many duplications of the same open-pollinated variety, which in several instances were only slight modifications. For instance, many known versions of the Reid Yellow Dent open-pollinated variety existed that bore different names, depending on individuals growing and practicing some mild selection within the variety.

Despite deficiencies in classification of germplasm in the USA, much is known concerning lineage of the germplasm and how the invaluable Corn Belt Dent race arose. Wellhausen et al. (1952) outlined the probable origin of the US maize germplasm (e.g., Hudson, 1994) that contributed to the Corn Belt Dent race (Fig. 11.1).

It is clear that the Corn Belt Dent varieties arose by repeated hybridization between the Northern Flints and the Southern Dents. Origin of the Northern Flints is still unclear and the Southern Dents somewhat conjectural. Brown and Anderson (1947) gave a detailed account of the Northern Flint racial complex, described some of the flint varieties, and listed the traits (hard kernels, low row number, cylindrical ears, and early maturity) that contributed to the Corn Belt Dent varieties.

![Fig. 11.1 Suggested probable origin of the Corn Belt Dents of the USA (Wellhausen et al., 1952)](image-url)
It has been suggested that the Northern Flints were derived from the Mexican race Harinoso de Ocho (Galinat and Gunnerson, 1963) derived from the southwestern USA (Mangelsdorf and Reeves, 1939) and from the San Marçenô and Serrano races of the highlands of Guatemala (Brown and Anderson, 1947). As Brown and Goodman (1977) pointed out, each of these conjectures has some faults. More evidence is available on the ancestral origin of the Southern Dent race, the other half of the parentage of the Corn Belt Dent race. Wellhausen et al. (1952) have involved several of the Mexican races in the parentage of the Southern Dents (Fig. 11.1). Brown and Anderson (1948) studied the Southern Dent complex in some detail and concluded that it was almost certainly derived from certain Mexican varieties. Brown and Goodman (1977) and Goodman and Brown (1988) also concluded that many Southern Dent varieties are related to the dents of central Mexico and that they are simply northern counterparts of races still prevalent in Mexico. There is not total agreement of the possible Mexican races involved in the Southern Dent race, but the Tuxpeno race seems prominent. Contributions of Southern Dents to Corn Belt Dents include high number of kernel rows, tapering ears, softer textured kernels, and pointed kernels (Brown and Anderson, 1948). Other traits that Southern Dents could have contributed are prolific tendencies and greater frequency of genes for disease and insect resistance.

Of the nine racial complexes described by Brown and Goodman (1977) and Goodman and Brown (1988) for the USA, five (Great Plains Flints and Flours, Pima-Papago, Southwestern Semidents, Southwestern 12 Row, and Derived Southern Dents) have had little impact on US maize breeding. Derived Southern Dents seem to have arisen from Southern Dents hybridized with (perhaps) Southeastern Flints, Northern Flints, and Corn Belt Dents. Derived Southern Dents, however, tend to have greater prolificacy than Southern Dents and include several varieties that have played an important role in breeding programs in the southeastern USA and as sources of prolificacy in Corn Belt Dents. The Great Plains Flints and Flours racial complex includes several distinctive varieties. Apparently this race was derived from the hybridization of Northern Flints and varieties from southwestern USA. Collections of Great Plains Flints and Flours are available but it does not seem that they have made a significant contribution to germplasm of the Corn Belt.

Maize breeders in the US Corn Belt have been concerned about the genetic variability of their breeding materials and have attempted to maintain some semblance of diversity in their breeding populations. Rapid development of breeding programs emphasizing inbred line development for use in hybrids and rapid acceptance of hybrids by farmers of the Corn Belt made the transition of the inbred–hybrid concept from theory to practicality. Empirical evidence showed that the greatest expression of heterosis in hybrids was from the use of inbred lines derived from divergent sources (Hallauer 1999b; Hallauer and Carena, 2000). Initial samplings of plants for deriving inbred lines were from open-pollinated varieties, but this soon changed to pedigree selection in F₂ populations formed from inbred lines considered superior to the average or possessing certain traits to correct known weaknesses of otherwise desirable lines.
It became apparent during the latter part of the 1940s that some order was needed to maintain the heterotic pattern of hybrids produced from lines derived from recycling of previously used lines. Although no attempt was made to classify material, collection and storage of germplasm in the form of open-pollinated varieties was emphasized in the 1930s. The minutes of the North Central Corn Improvement Conferences frequently included reports by subcommittees on the preservation of germplasm. The extent of the preservation depended on the interests of the individuals within the 12 states in regard to collecting, maintaining, and describing the collections. Each state was encouraged to maintain storage of germplasm locally and at some central storage agency. Germplasm of open-pollinated varieties was collected for preservation, but the extent of its inclusion in breeding programs was usually minimal. Hybrid trials including US public lines from three different maturity zones were discontinued, unfortunately the last trial from the North Central Corn Breeding Research Committee (NCR167, now NCCC167) was from the early-maturing zone (FAO 100–300) in 2008 including Cornell and North Dakota State University lines crossed with industry testers.

Renewed interest in maize germplasm occurred after the B. maydis outbreak on T-cytoplasm in 1970. In 1975 maize breeders in the 12 north central states appointed a new subcommittee on genetic vulnerability that was charged with the responsibility to assess the stage of germplasm base in the North Central Region of the USA. The subcommittee surveyed all the public maize breeders in the North Central Region to determine the scope of the germplasm base of their breeding programs. A summary of the subcommittee’s report was included in the minutes of the US North Central Corn Breeding Research Committee (NCR-2) Report (1977) (Tables 11.2, 11.3, and 11.4) and a brief description of each population was given.

Surprisingly, the germplasm base of US Corn Belt breeding programs was not as restricted as originally feared; 246 populations were currently undergoing some form of selection, and additional 200 populations included in the breeders’ inventory were in storage and available for use. There were several duplications of basic germplasm undergoing selection (Table 11.2). The most striking example was the 24 versions of Iowa Stiff Stalk Synthetic being used.

Improvement for grain yield was the most common single trait under selection (Table 11.3); disease resistance also had a high priority. The primary traits of selection are listed in Table 11.3 and usually more than one trait was included, particularly in selection programs for yield. For instance, selections based on mechanically harvestable yield would include selection for stalk quality (disease resistance) and ear droppage (corn borer resistance).

Mass selection was the most commonly used breeding method (Table 11.4). In 36 instances S_1 progeny evaluation was used, with about equal use of the other methods. Of the populations actively under selection, 62 (25.2%) included germplasm that was considered exotic compared to 54.1% of the populations being synthesized from Corn Belt lines (excluding Iowa Stiff Stalk Synthetic) while 10.8% were open-pollinated varieties. The inactive populations were nearly equal for the US Corn Belt open-pollinated varieties (20.6%) and populations that included exotic germplasm (19.5%).
Table 11.2 Populations in the US North Central Region survey classified by origin and whether they are currently undergoing selection

<table>
<thead>
<tr>
<th>Origin</th>
<th>Active</th>
<th></th>
<th>Inactive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Krug</td>
<td>9</td>
<td>3.6</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Lancaster</td>
<td>4</td>
<td>1.6</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>Hays Golden</td>
<td>7</td>
<td>2.8</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Other open-pollinated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varieties</td>
<td>7</td>
<td>2.8</td>
<td>38</td>
<td>19.0</td>
</tr>
<tr>
<td>Exoticsb</td>
<td>62</td>
<td>25.2</td>
<td>39</td>
<td>19.5</td>
</tr>
<tr>
<td>Stiff Stalk Synthetic</td>
<td>24</td>
<td>9.8</td>
<td>8</td>
<td>4.0</td>
</tr>
<tr>
<td>Other synthetic varieties</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-linec</td>
<td>2</td>
<td>0.8</td>
<td>10</td>
<td>5.0</td>
</tr>
<tr>
<td>Early (A.E.S. 400 – 600)</td>
<td>42</td>
<td>17.1</td>
<td>45</td>
<td>22.5</td>
</tr>
<tr>
<td>Late (A.E.S. 700 – 800)</td>
<td>89</td>
<td>36.2</td>
<td>56</td>
<td>28.0</td>
</tr>
<tr>
<td>Total</td>
<td>246</td>
<td></td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

aReport of the North Central Corn Breeding Research Committee (NCR-2) (1977, p. 66)
bPopulations that include at least some exotic germplasm
cPopulations that include different versions of the same inbred line

Table 11.3 Traits undergoing selection in the populations included in the US North Central Region survey

<table>
<thead>
<tr>
<th>Trait</th>
<th>Number of populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain yield</td>
<td>44</td>
</tr>
<tr>
<td>Grain yield and maturity</td>
<td>9</td>
</tr>
<tr>
<td>Agronomic traits</td>
<td>43</td>
</tr>
<tr>
<td>Insect resistance</td>
<td>14</td>
</tr>
<tr>
<td>Disease resistance</td>
<td>57</td>
</tr>
<tr>
<td>Chemical composition</td>
<td>28</td>
</tr>
<tr>
<td>Adaptation</td>
<td>15</td>
</tr>
<tr>
<td>Unidentified</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>246</td>
</tr>
</tbody>
</table>

aReport of the North Central Corn Breeding Research Committee (NCR-2) (1977, p. 67)

The survey of germplasm included in the US Corn Belt breeding programs indicated that the problem of genetic vulnerability was less serious than previously assumed. At the time of the survey in the late 1970s, a wide range of germplasm was undergoing selection for several different traits by the use of several different selection procedures. The situation seems to be different 30 years later with very few public programs conducting germplasm improvement and cultivar development and consolidation of the seed industry. Even if the genetic diversity in reserve was still large, how could one reconcile the evidence that there was a dependence on only a few inbred lines for use in hybrids (Corn, 1972, p. 105; Zuber, 1975; Mikel, 2006,
Table 11.4 Selection methods being used for the populations included in the US North Central Region survey

<table>
<thead>
<tr>
<th>Method of selection</th>
<th>Number of populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>140</td>
</tr>
<tr>
<td>S₁ per se</td>
<td>36</td>
</tr>
<tr>
<td>Full-sib plus mass</td>
<td>11</td>
</tr>
<tr>
<td>S₂ per se</td>
<td>11</td>
</tr>
<tr>
<td>Inbred tester</td>
<td>11</td>
</tr>
<tr>
<td>Reciprocal full-sib</td>
<td>8</td>
</tr>
<tr>
<td>Full-sib</td>
<td>7</td>
</tr>
<tr>
<td>Reciprocal recurrent</td>
<td>4</td>
</tr>
<tr>
<td>Modified ear-to-row</td>
<td>4</td>
</tr>
<tr>
<td>Half-sib</td>
<td>3</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>246</td>
</tr>
</tbody>
</table>

*Report of the North Central Corn Breeding Research Committee (NCR-2) (1977, p. 68)

2008; Mikel and Dudley, 2006) and in the range of germplasm included in breeding programs? In all instances the surveys were for lines released from publicly supported breeding programs. It seems reasonable to assume that a similar survey involving privately supported breeding programs would have given similar results (Duvick 1975, 1992). It does not seem that breeding programs are extremely limited in genetic variability, but the economic restraints demanded by modern farmers guarantee that genetic variability within and among fields of maize will be much less than when open-pollinated varieties were predominantly used. Modern farmers demand the use of hybrids that have proven yield performance and uniform plant type and maturity. A unique combination of two inbred lines that provides the desired hybrid will be extensively used. The hybrid seed industry is very competitive and will provide seed to meet the demands. In many cases, the same or similar lines will be used (Mikel, 2008) especially among hybrids carrying different transgenic events. Extensive use of a few inbred lines of similar genetic background in the production of hybrids is a cause for concern of genetic vulnerability. Maize breeders are cognizant of the problem but they may not be able to do much about it. However, if strong public programs are still active integrating pre-breeding (e.g., germplasm improvement) with inbred line development (Carena, 2008b) they may have the materials available if and when the need arises.

11.6 Germplasm Improvement

Collection, classification, and maintenance of germplasm have been emphasized to maintain its availability for future use. Although original forms of races and varieties of maize are fast disappearing because of the change to hybrids, preservation of race collections is less than adequate because of limited interest, funds, and facilities. Races and varieties are populations of individuals, each with unique genotypes that
require sizable (e.g., 200 – 500) numbers and samples (e.g., a large ear number) to maintain genetic characteristics. Thousands of collections have been made, but the potential and future use of the materials is limited if the collections are not properly stored and maintained in the short and long term.

Three major centers were established for the storage and preservation of collections in the Western Hemisphere:

1. Maize Germplasm Bank at the Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), Chapingo, Mexico, which serves Central America and the Caribbean area.

2. Seed Storage Center at Medellín, Colombia, which serves Colombia and the other Andean countries (Bolivia, Chile, Ecuador, Peru, and Venezuela). It is maintained jointly by the Colombian Ministry of Agriculture and the Rockefeller Foundation.

3. Brazilian Germplasm Bank at CNPMS-EMBRAPA (Sete Lagoas, M.G.) (formerly at Instituto de Genética, Escola Superior de Agricultura ‘Luiz de Queiroz,’ Piracicaba, Brazil). This center serves Argentina, Brazil, eastern Bolivia, Paraguay, Uruguay, and Guyana.

Additional collections are maintained at INIA, Peru la Molina; Argentina-INIA; USDA National Seed Storage Laboratory at Fort Collins, Colorado; and the USDA North Central Regional Plant Introduction Station at Ames, Iowa. Many of the collections at the different storage and preservation centers are duplicates, and the number of collections available probably exceeds those for most other crop species. Lonnquist (1974) indicated there are over 25,000 collections at the storage centers in the Western Hemisphere and about 6,000 collections in the European–Mediterranean area.

Most breeding projects also maintain a limited number of collections, which some are included in the seed storage centers. Duplications of collections in the storage centers and breeding projects are desirable because of differences that can arise from their maintenance from genetic drift. However, most unreleased genetic materials depend on local cold storage units often dependent on breeder’s budgets and under risk. In fact, land-grant Institutions often lack plans to preserve germplasm from programs that are discontinued. A national program (e.g., USDA-CAP) with the priority of long-term cold storage preservation of maize improved germplasm would be desirable. Individual state cold storage units could risk the availability of improved and locally adapted germplasm.

Paterniani and Goodman (1977) discussed problems associated with preservation of collections such as labor and facilities required to maintain the collections; population sizes required to minimize the effects of inbreeding, loss of genes, and genetic drift; contamination; amount of seed required to meet the demands of researchers; and loss of materials that are poorly adapted to the area used for propagation. All problems associated with preservation of germplasm are related to the one important problem of population size.

Duplicate samples at storage centers probably become quite different for some traits because of the limited population sizes used in their maintenance. To overcome
some of these problems, Paterniani and Goodman (1977) used a more practical approach for preservation of their collections by compositing similar original samples. Use of this method provides more adequate maintenance of a smaller representative number of populations.

Maize populations are usually maintained by hand sib-pollinations that require extensive supplies and labor, and only a limited number of plants may be included for each population. Hence the problems of expenses, genetic drift, seed supplies, and contamination can arise. Omolo and Russell (1971) propagated by hand-controlled pollinations 500, 200, 80, 32, and 18 plant populations for five successive generations of the open-pollinated variety Krug. They wanted to determine what size of population was required in order to maintain itself without causing significant genetic changes. Significant yield decrease was obtained as sample size decreased, and they concluded that a sample of 200 plants was adequate to maintain a heterogeneous population by hand sib-pollination. Only if some inbreeding could be tolerated and reproduction was infrequent would an 80-plant sample be adequate. The results of Omolo and Russell (1971) emphasize that adequate population sizes are important to retain the original genetic variability, but the size also imposes serious problems in preservation of germplasm. Reproduction of the collections in isolation fields would be preferable, but locating sites, field husbandry, and adequate isolation (at least 200 m) make this approach unfeasible. Adequate isolation would minimize all problems associated with germplasm maintenance except for cost of labor and facilities.

The effective size of a population is an important aspect of germplasm preservation (see Chapter 9). The level of maintenance of genetic properties of a population depends partly on the number $N$ of seeds or individuals in the population (census number) and primarily on the number of individuals intercrossed in previous generations, which reflects the effective number $N_e$. The effective number depends also on how female and male gametes are sampled to originate the descendant population. In addition, if the population is treated as a dioecious species, $N_e$ will depend on the number of males and females taken for crosses. If the population is crossed as a monoecious species each plant is potentially assumed to participate as a source of male and female gametes.

The largest possible effective size of a population occurs in homozygous self-pollinated species when one seed or an equal-sized sample is taken from each plant to originate the next generation. In cross-pollinated species, such as maize, effective size depends on the crossing system and the manner in which female and male gametes are sampled. Therefore, maintenance of germplasm stocks can be performed by any of the following procedures:

1. Population is treated as a monoecious species. Then each plant in the population is a potential source of female and male gametes, which can be sampled with or without control on their relative number. Three sampling procedures are possible:
(a) Control of the number of female and male gametes, which is only possible under controlled hand pollination so that each male plant contributes an equal number of gametes to the next generation. To obtain equal contributions of male gametes an equal number of seeds must be taken from each pollinated ear, thus resulting also in equal numbers of female gametes contributed to the next generation.

(b) Control of the number of female gametes only. An equal number of contributed female gametes is obtained by taking an equal number of seeds from each ear (female parent), which is pollinated at random so that each male plant contributes unequally to the next generation.

(c) No control of either female or male gametes. No control results when pollination is at random (no control of the number of male gametes) and samples of varied sizes are taken from each ear (no control of female gametes). This procedure is common when all the ears of an open-pollinated field are harvested together (by hand or machine) and one sample is taken from the bulked seeds.

2. Population is treated as a dioecious species. A monoecious species like maize can be crossed as a dioecious species, provided that some plants are used as male and others as female parents. Separation of male and female plants in a field of maize can be accomplished through controlled hand pollination or by detasseling plants or plant rows. Control of the number of female and male gametes results in the following cases:

(a) Control of both female and male gametes, which requires either controlled hand pollination in plant-to-plant (male × female) crosses or one plant used as a male to pollinate several plants as females. In the first case we have the same number of males $N_m$ and females $N_f$. For instance, $N_m = N_f = N/2$, where $N$ is the total number of plants. In the second case the number of females will be greater than the number of males. For instance, if one male is used to pollinate four female plants, then $N_m = N/5$ and $N_f = 4N/5$. In either case, equal-sized samples must be taken from each pollinated ear to control the number of female and male gametes.

(b) Control of only the number of female gametes, which means that pollination is completely at random so that each male plant contributes unequally to the next generation. The easiest way to do this is by detasseling a proportion of rows in a field of maize; e.g., by detasseling alternate rows we have $N_m = N_f = N/2$, assuming a constant number of plants per row. To have control of the number of female gametes, samples of equal size must be taken from each pollinated ear in the female rows.

(c) No control of either female or male gametes, which requires either hand pollination or random pollination in a detasseled field with unequal-sized samples from each pollinated ear. This would result when a single sample of seeds is taken from a bulk of all ears from female plants.
Effective sizes for populations resulting from each of the given procedures are shown in Table 11.5.

For calculation of the effective size of a population there are two different approaches: inbreeding effective size and variance effective size procedures. They are equivalent for a constant population size over generations.

With control of the number of female and male gametes when the population is crossed as a monoecious species [case 1(a)], the effective number can double in relation to the simplest procedure where there is no control on gametes of either sex [case 1(c)]. This is particularly important for small populations, which occur very frequently with introduced germplasm. On the other hand, seed sampling and controlled hand pollinations are not difficult in the maintenance of relatively small populations. When the population is pollinated as a dioecious species, the effective number is the same as when treated as a monoecious species if the same number of male and female parents are used. Otherwise, the effective number tends to decrease in the direction of the sex that participates with the smaller number, because $N_e$ is proportional to the harmonic mean of $N_m$ and $N_f$ and the harmonic mean is more strongly influenced by the smaller values (Crow and Kimura, 1970). Therefore, for $N_f > N_m$, the number of males is more important than the number of females in determining the effective size of the population and vice versa.

When considering several generations in a random mating population, the effective number may change with differences in sample sizes and procedures. In a random mating population where size fluctuates, it would be desirable to know what population of constant size would give the same effective number. Crow and Kimura (1970) showed that the effective population number is roughly the harmonic mean of the various values. Therefore $\left(1/N_e\right) = \left(1/t\right) \sum_i \left(1/N_i\right)$, where $t$ is the number of generations considered. As an example, consider four generations where the effective numbers are 100, 100, 20, and 1000. The effective number of a corresponding

<table>
<thead>
<tr>
<th>Crossing systemb</th>
<th>Gamete control Female</th>
<th>No. of parents Female</th>
<th>Effective number ($N_e^c$)</th>
<th>Relative numberd</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Monoecious</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a)</td>
<td>With</td>
<td>$N$</td>
<td>$2N$</td>
<td>200</td>
</tr>
<tr>
<td>(b)</td>
<td>With</td>
<td>$N$</td>
<td>$4N/3$</td>
<td>133</td>
</tr>
<tr>
<td>(c)</td>
<td>No</td>
<td>$N$</td>
<td>$N$</td>
<td>100</td>
</tr>
<tr>
<td>2. Dioecious</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a)</td>
<td>With</td>
<td>$N_f$</td>
<td>$8N_mN_f/(N_m+N_f)$</td>
<td>200</td>
</tr>
<tr>
<td>(b)</td>
<td>With</td>
<td>$N_f$</td>
<td>$16N_mN_f/(3N_m+N_f)$</td>
<td>133</td>
</tr>
<tr>
<td>(c)</td>
<td>No</td>
<td>$N_f$</td>
<td>$4N_mN_f/(N_m+N_f)$</td>
<td>100</td>
</tr>
</tbody>
</table>

---

aThe authors are indebted to R. Vencovsky for derivation of the formulas
bSee text for explanation
cIndicates constant population number of each procedure, without random elimination of plants
dRelative to 1c ($N_e = N$) when monoecious and to 2c ($N_m = N_f$) when dioecious

<table>
<thead>
<tr>
<th>Crossing system</th>
<th>Gamete control</th>
<th>No. of parents</th>
<th>Effective number ($N_e$)</th>
<th>Relative number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Monoecious</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a)</td>
<td>With</td>
<td>$N$</td>
<td>$2N$</td>
<td>200</td>
</tr>
<tr>
<td>(b)</td>
<td>With</td>
<td>$N$</td>
<td>$4N/3$</td>
<td>133</td>
</tr>
<tr>
<td>(c)</td>
<td>No</td>
<td>$N$</td>
<td>$N$</td>
<td>100</td>
</tr>
<tr>
<td>2. Dioecious</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a)</td>
<td>With</td>
<td>$N_f$</td>
<td>$8N_mN_f/(N_m+N_f)$</td>
<td>200</td>
</tr>
<tr>
<td>(b)</td>
<td>With</td>
<td>$N_f$</td>
<td>$16N_mN_f/(3N_m+N_f)$</td>
<td>133</td>
</tr>
<tr>
<td>(c)</td>
<td>No</td>
<td>$N_f$</td>
<td>$4N_mN_f/(N_m+N_f)$</td>
<td>100</td>
</tr>
</tbody>
</table>
population of constant size would be $N_e = 56.3$. If one wants to compensate for the bottleneck in the third generation and sample 10,000 individuals instead of 1,000 in the fourth generation, little is gained; the effective population size would correspond to a population of constant size $N_e = 57.1$. This example shows that in maintaining germplasms care must be taken to keep the population size at least nearly constant over generations, because the decrease in population size in one generation is not easily recoverable in the next.

When there is a reduction in sample size in one generation after some generations of constant size, it is possible to recover the original trend of effective size by using one or more generations of larger sample sizes. Over $t$ generations we have $t - k - 1$ generations of constant size ($N_1$) followed by one generation of reduced size ($N_2 < N_1$). Then, we want to recover the original trend ($N_1$) of effective size in the next $k$ generations with effective size of $N_3$. It can be shown that this is possible only for $(k + 1)N_2 > N_1$ and that the required effective size in the last $k$ generations must be $N_3 = (kN_1 N_2) / [(k + 1)(N_2 - N_1)]$. Note that $N_3$ does not depend on the number of initial generations under constant size.

To recover the loss in effective size in just one generation after the bottleneck, then $k = 1$ and it is necessary that $N_2 > N_1/2$. In the example given, we had $N_1 = 100$, $N_2 = 20$, and $N_3 = 1000$. In this case we have $N_2 < N_1/2$, showing that it is not possible to recover in one generation the loss in the third generation. However, if $N_2 = 60$, it would be possible to recover the loss by taking $N_3 = 300$; i.e., effective sizes of 100, 100, 60, and 300 in four generations would be equivalent to a constant population size $N_e = 100$. For $N_2 = 50$, recovery would be possible by taking more than one generation of larger sample sizes. Thus, effective sizes of 100, 100, 50, 200, and 200 would give an effective population size equivalent to a population of constant size $N_e = 100$.

Race collections are irreplaceable and represent a source of germplasm that is either limited or not available at the present. Races of maize have evolved over hundreds of years of natural and artificial selection that developed unique populations adaptable to a wide range of environmental conditions. Seed storage banks are a repository of genes and gene combinations that otherwise would have been lost because of use of hybrids, increased cultivation of land, and decrease in numbers of people involved in agriculture. Genetic improvement of hybrids has been made from use of only a limited amount of the total germplasm available (e.g., Mikel, 2008). Brown (1975) has stated that in the USA more than 90% of the breeding effort is devoted to germplasm whose origin is traced to not more than 3 of the 130 existing races. Hence, he continues, US maize improvement programs have largely ignored 98% of the germplasm that makes up $Z. mays$. Galinat (1974) also has expressed concern about the genetic erosion of maize and the loss of natural genetic variation because of the pressures for developing uniform varieties that produce the greatest yields. In the future, seed repositories would become immediately useful in breeding programs for increasing the genetic variability of the materials undergoing selection. The only recourse for recapturing natural variation is to take advantage of the seed storage centers. They contain material that can be incorporated into long-range breeding programs. They are sources of genes for resistance
to maize pests (to meet emergencies); sources of germplasm to further basic studies on heterosis, relationships, and origin of maize; and a wealth of genes and combinations of genes for basic genetic studies. The importance of indigenous maize germplasm is emphasized by all the authors listed in Table 11.1. It is imperative that efforts be expended to maintain and utilize the reservoir of maize germplasm for future breeding programs.

11.7 Potential and Use of Exotic Germplasm

Double-cycle breeding and the limited utilization of maize genetic diversity could limit future significant genetic gains. The possibilities of including exotic germplasm in maize breeding programs have been emphasized by reports listed in Table 11.1 as well as by several research groups (Brown, 1953, 1975; Griffing and Lindstrom, 1954; Wellhausen, 1956, 1965; Rinke and Sentz, 1961; Leng et al., 1962; Paterniani, 1962; Goodman, 1965b; Brandolini, 1969; Lonnquist, 1974; Brown and Goodman, 1977; Albrecht and Dudley, 1987; Goodman and Brown, 1988; Iglesias and Hallauer, 2003; Pollak, 2003; Menz and Hallauer, 1997; Salhuana et al., 1998; Pollak, 2003; Carena and Wicks III, 2006; Carena, 2008a, Hallauer and Carena, 2009; Carena et al., 2009a, etc.). In most instances, it was also emphasized that the immediate usefulness of exotic germplasm may be limited to single-gene transfers but larger percentages of exotic germplasm have already been utilized and incorporated in breeding programs with successful results. The wealth of germplasm available is staggering if one considers the volume of the maize collections that have been assembled for possible use (Lonnquist, 1974).

Exotic germplasm can have several connotations. For applied breeding programs exotic germplasm includes genetic material (which does not need to be foreign) that is not presently adapted to local maize breeding programs and, consequently, does not have immediate usefulness without selection for adaptation and needs improvement before successful utilization (Carena and Hallauer, 2001). The small germplasm sample being utilized in the US Corn Belt today is a consequence of insufficient germplasm evaluation and further improvement. US domestic germplasm has received less attention (Kauffman et al., 1982; Hallauer et al., 1988), even though it has several advantages over tropical germplasm. Because of a similar evolutionary history to elite germplasm, US domestic germplasm will require less time for adaptation than other exotic germplasm sources. Desirable traits are also not masked by photoperiod responses as in the case of tropical material.

Paterniani (1962) distinguished two possible alternatives in exotic germplasm: (1) races or varieties having a broad genetic base and (2) inbred lines having a narrow base. Races and varieties can be used either in population improvement programs or in inter-varietal crosses to determine the hybrid vigor expressed in crosses. Inbred lines can be used either for conventional hybrids or to synthesize populations for special purposes. The choice between the two alternatives and its success depends on the level of maize improvement, the social and economic situation, and genetic variability and potential of the local germplasm.
In spite of the extreme variability within *Z. mays* the gross chromosome morphology is rather uniform among the many races, varieties, and strains (Galinat, 1977). Similar chromosome numbers and morphology of exotic and adapted germplasm permit wide crossing. However, the immediate effects of crossing exotic germplasm with native germplasm and initiating selection are often disappointing. Productivity and desirable segregates are usually limiting and the material is usually discarded as not being promising. To obtain gene combinations that have efficient biochemical functions within the genome, it is necessary to allow genetic recombination accompanied by mild selection. Lonnquist (1974) emphasized the necessity and importance of recombination and the possible problems of linkage. The only recourse for the use of exotic germplasm in, say, the US Corn Belt is patience and adequate recombination after several generations of random mating with mild selection pressure (Brown, 1953; Lonnquist, 1974).

Troyer and Brown (1972) demonstrated the effectiveness of gradual introductions of exotic germplasm into adapted germplasm of the US Corn Belt. They crossed Mexican germplasm with US Corn Belt lines and grew the materials in isolated fields for 10 years to allow for recombination. Mass selection for recombinant genotypes that had desirable plant traits was carried out before intense selection was practiced for earliness. Other successful and cost-efficient mass selection methods to adapt exotic germplasm were discussed in previous chapters.

Each race of germplasm has evolved over a period of time to develop a population of genotypes that are physiologically adapted to a particular ecological niche. The collection of genetic factors for each race has been assembled over time to develop genotypes that function efficiently to survive and propagate the race. Although the Corn Belt Dent race has not been subjected to selection for as long as most of the other races described in Table 11.1, it has been under relatively intense selection pressure for the factors that collectively formed the highly productive race. The crossing of races would disrupt the harmonious gene combinations of each race. It would be similar to the situation that existed in the past when a race was established by crossing two existing races. An example from Wellhausen et al. (1952) shows that the Harinoso Flexible and Teocintle races were crossed with form the race Olotillo; the Harinoso de Guatemala and Teocintle races were crossed to form the race Tepecintle; finally the Olotillo and Tepecintle races were hybridized to form the important race Tuxpeno. Tuxpeno is one of the more productive and desirable modern races of Mexico and has been a source of germplasm for the Southern Dents of the USA. The Corn Belt Dent race, in turn, was formed by the hybridization of the Southern Dent and Northern Flint races. Successive hybridizations were followed by a period of intermating and selection to gradually evolve into the new races. Interracial hybridization seems to have been an important factor in the development of superior germplasm (Wellhausen et al., 1952).

The evolution of the formation of new races would be similar to the methods needed for incorporating exotic germplasm into adapted germplasm. The key issue seems to be allowing sufficient time for the integration of the genetic factors from the two sources of germplasm. If the evolution of the development of the superior new races is to be mimicked in a breeding program, it seems imperative to permit
adequate recombination with only mild selection pressure to sort out the desirable alleles in the new breeding population. After the derived germplasm has developed to the point of seeming adaptability to the particular environment, more intense selection accompanied by inbreeding may be initiated.

Hallauer and Carena (2009) summarized information available from the use of exotic germplasm in the USA under the following categories: variability among and within exotic germplasm, heterosis expressed among exotic varieties and among exotic × adapted varieties, effects of selection within exotic germplasm and within populations formed by crossing exotic and adapted varieties, and potential of exotic germplasm as sources for line development and genes for disease and insect resistance. Although Wellhausen et al. (1952), Brown (1953), Wellhausen (1956, 1965), Leng et al. (1962), Goodman (1985, 1999b), and others have emphasized the importance and potential of the germplasm from outside the USA, the amount of information from the use of this material is rather limited. Results reported generally are positive, but they may be misleading as an indication of the effort that has been expanded on screening and evaluating exotic germplasm. It probably has been included in most applied breeding programs at some time and the results were either negative or not reported. The survey of the North Central Region of the USA shows that 25.2% of the populations actively undergoing selection included at least some exotic germplasm (Table 11.2). The proportion of exotic germplasm in commercial applied breeding programs may be even greater.

Goodman (1965b) reported data that include a critical comparison of the estimates of genetic variability for an adapted population (Corn Belt Composite) and a population that included exotic germplasm (West Indian Composite). His results showed that genetic variability was greater for the population that included exotic germplasm and not at the expense of lower yields. Predicted gain was greater in the West Indian Composite and the opportunities seemed good that material developed from the West Indian Composite would contribute to the heterosis of hybrids, which was substantiated by Eberhart (1971) in a diallel series of variety crosses that included West Indian Composite. Shauman (1971) also reported greater estimates of additive genetic variance in the hybrid population of Krug × Taboncillo 13 Hi Synthetic 3 than in the adapted variety Krug.

Heterosis among exotic varieties and among exotic × adapted varieties has been determined for several races and varieties (Wellhausen, 1956, 1965; Vasal et al., 1999, see Chapter 10). Evaluations of races and varieties (both adapted and exotic) have not been as extensive as desirable, but during the past 25 years the Latin American Maize Project (LAMP) and the Germplasm Enhancement of Maize (GEM) initiatives were established to alleviate the previous limitations of programs conducted on a limited scale by individuals (Pollak 2003). Widespread trials of races and varieties would assist in the identification of those most desirable for breeding programs with added information on the manifestation of heterosis.

Wellhausen et al. (1952) and others (Table 11.1) have emphasized that heterosis information would be a valuable additional aid in classification of maize germplasm. Heterosis was usually observed for crosses that included exotic germplasm (Oyervides-Garcia et al., 1985; Vasal et al., 1992a, 1992b, 1999; Michelini and
Hallauer, 1993; Eschandi and Hallauer, 1996). The contribution of genetic diversity to manifestation of heterosis, however, seems to have a limit. Moll et al. (1962, 1965) in studies of crosses among adapted and exotic varieties found that heterosis increased as presumed genetic diversity increased but decreased in variety crosses that were assumed to be most genetically diverse. The possible explanation is that the combination of genetic factors from extremely diverse germplasm was too great to allow compatible functioning of the physiological mechanisms. Before extremely diverse exotic germplasm is introduced it would be helpful to know these relationships in order to increase efficiency of a breeding program. The problem may be alleviated to some extent because selection for adaptability of exotic germplasm is necessary before it is useful in a breeding program (Hallauer, 1999a).

Population improvement programs conducted in exotic and semi-exotic populations usually have been effective in the limited instances reported (Pandey and Gardner, 1992; Hallauer, 1992, 1999a). Selection for improvement was usually conducted in populations formed by crossing adapted and exotic germplasm, but Hallauer (1999a) selected directly in ETO Composite, Antigua Composite, Tuxpeno Composite, Suwan-1, and Tuson Composite for earlier maturity; after six to eight cycles of phenotypic (mass) selection for earlier silking ears, the 100% tropical populations were adapted to temperate environments. Three cycles of stratified mass selection have been sufficient to move US Corn Belt temperate populations to North Dakota environments (Eno and Carena, 2008) while five to six cycles have been enough to adapt highland tropical populations to the same environments. Initial selection generally is for adaptability, i.e., maturity and reduced plant stature. Mass selection has proven effective for adaptability (Hallauer and Sears, 1972; Troyer and Brown, 1972; Hallauer, 1999a; Carena et al., 2008; Hallauer and Carena, 2009) in spite of the complexity reported in the genetic architecture of maize flowering (Buckler et al., 2009). Salvi et al. (2002) mapped a quantitative trait loci (QTL) controlling the transition from the vegetative to the reproductive phase in maize (Vgt1) on chromosome 8 based on a mapping population derived from the cross N28 × C22–4. The strain C22–4 is nearly isogenic to N28 with an introgression from the early maize variety Gaspé Flint. As a consequence, Vgt1 could be used in a marker-assisted selection (MAS) program to produce maize cultivars that are 1 week earlier than the original version. It is recognized the amount of efforts and time invested in isolating this gene as well as its potential usefulness in basic science as a genetic system. However, since early flowering is a highly heritable trait and easy to measure, alternative and cheaper selection methods are available if selection and improvement is the priority of the research. Heritability is rather high for adaptability traits and mass selection techniques permit additional recombination with a mild selection pressure. Number of cycles of mass selection necessary to attain acceptable standards of adaptability depend on the germplasm included, selection intensity, and the range in latitude over which the germplasm is being transferred. Recurrent selection in populations that include exotic germplasm was initiated for yield improvement to compare rates of gain and genetic variability; S2 progenies have been evaluated in replicated field trials (Hallauer, 1978). Populations undergoing S2 recurrent selection included different proportions of
exotic germplasm: BS16 was developed by mass selection for early maturity from ETO Composite; BS2 was formed by crossing ETO Composite with six early inbred lines followed by five generations of intercrossing; BSTL was developed by crossing Lancaster Surecrop with Tuxpeno and backcrossing to Lancaster Surecrop; and Krug Hi I Synthetic 3 was an adapted Corn Belt variety. The relative proportions of exotic germplasm are 100, 50, 25, and 0% for BS16, BS2, BSTL, and Krug Hi I Synthetic 3, respectively. One objective of the four concurrent selection programs was to determine what effect if any the different proportions of exotic germplasm have on gain and variability in successive cycles of selection. Preliminary results did not show any striking differences among the four populations (Hallauer, 1978), but critical comparisons will not be available until after additional cycles of selection. Wellhausen (1965) was of the opinion that only small doses (25% or less) of exotic germplasm should be incorporated initially into adapted populations. Following the suggestion of Wellhausen, Whitehead et al. (2006) conducted an extensive program to integrate elite, selected germplasm developed by CIMMYT with elite, selected US Corn Belt germplasm in Iowa. Heterotic groups for the respective areas were considered in making the germplasm crosses: subtropical maturity – BSSS crossed with Tuxpeno and non-BSSS crossed with non-Tuxpeno; and tropical maturity – BSSS crossed with Tuxpeno and non-BSSS crossed with non-Tuxpeno. Crosses and backcrosses to temperate populations were made in Mexico with subsequent testing of backcross progenies and their testcrosses conducted in US Corn Belt. Based on backcross and testcross trials, selected backcross progenies were intermated to form four populations: two subtropical × US Corn Belt populations of BSSS × Tuxpeno (BS35), non-BSSS × non-Tuxpeno (BS36), two tropical × US Corn Belt populations of BSSS × Tuxpeno (BS37), and non-BSSS × non-Tuxpeno (BS38) (Hallauer, 2005). Data are not available to determine if the germplasm developed by Whitehead et al. (2006) with 25% tropical germplasm has any advantages, or disadvantages, compared with the germplasm sources (100% tropical) phenotypically selected for adaptation to temperate environments (Hallauer, 1999a). Further pre-breeding, however, is needed in all instances. Pre-breeding is not a recent concept and has been an important component in the development of single-cross hybrids. Pre-breeding is the long-term conservation and utilization of genetic resources linked to an efficient cultivar development process (Carena, 2008b). It includes the introduction, adaptation, evaluation, and improvement of germplasm resources for use in breeding programs (Hallauer and Carena, 2009). Pre-breeding develops germplasm resources that are either directly or indirectly used to develop new cultivars (e.g., recurrent selection linked to pedigree selection).

Evaluations of materials that included exotic germplasm have been reported for development of lines for use in hybrids (Griffing and Lindstrom, 1954; Paterniani, 1964; Efron and Everett, 1969; Nelson, 1972; Goodman 1999a; Carena et al., 2009a,b), sources of disease (Kramer and Ullstrup, 1959) and insect (Sullivan et al., 1974; Carena and Glogoza, 2004) resistance, and silage production (Thompson, 1968). Nelson (1972) has used exotic germplasm in an applied breeding program developing lines for use in hybrids in the southern part of the USA. He has developed lines that have played a prominent role in the production of hybrids grown in
that area. For the northern USA, early × late crosses and backcrosses have been the strategies to move 0, 25, and 50% exotic germplasm northward and westward and develop new and unique early-maturing lines (Rinke and Sentz, 1961; Hallauer et al., 1988; Hallauer and Carena, 2009; Carena et al., 2009a,b). Other maize breeders, public or private, undoubtedly have integrated exotic germplasm into their breeding programs; but the stage of development either does not permit their use in hybrids or the material is used in proprietary hybrids and the extent of use is unknown. Considerable screening for sources of resistance to diseases and insects have occurred in breeding programs and a significant amount of these materials included some proportion of exotic germplasm. In Table 11.2, populations that include at least some exotic germplasm were one of the largest categories currently undergoing active (25.2%) or inactive (19.5%) selection, and the number undergoing selection for disease (57%) or insect (14%) resistance was substantial (Table 11.3). Considerable effort was being expended on populations that included some exotic germplasm, but in most instances results were in the preliminary or initial stages of selection and/or were discontinued.

Wellhausen et al. (1952), Brown (1953), Wellhausen (1956, 1965), and Leng et al. (1962) have emphasized the importance of exotic germplasm and thought US maize breeders were in an excellent position to make use of exotic germplasm. A concerted effort to exploit the wealth of germplasm available for use was indirectly proposed (Brown, 1975). Maize breeders in the USA have been making significant genetic improvements in their hybrids through use of conventional breeding procedures and adapted germplasm (Russell, 1974; Duvick, 1977, 1992; Duvick et al., 2004). The status quo has been satisfactory on a short-term basis while incorporation of exotic germplasm into adapted populations or its direct use would be included in long-range objectives of the breeding program. Exotic germplasm must include useful genes, but they will not be available until they are incorporated with the highly productive adapted germplasm. It will require time and patience. Immediate payoffs are not to be expected but long-range payoffs seem likely. Most of the evidence reported from use of exotic with adapted germplasm has been encouraging. The initiation of selfing in recently hybridized exotic by adapted germplasm usually has been negative as inbreeding depression is severe and few vigorous lines are obtained. Frequently the material was discarded and the use of exotic germplasm seemed fruitless. A common error was not recombining the best progenies and initiating another cycle of recombination of the best material. Additional selection and recombination would permit further integration of linkage blocks and choice of genes desired for the particular environment.

Choice of exotic germplasm to include has to be considered. Wellhausen (1965), Brown and Goodman (1977), Goodman and Brown (1988), Vasal (1999), Goodman (1999b), and others have discussed some of the Mexican and Caribbean races and populations that seem promising. Lack of systematic information on the relative merits of different races in breeding programs prevents making precise recommendations for given situations; this was emphasized by the authors in Table 11.1. Hence, choice of exotic germplasm often depended on limited experience in the area of adaptation and limited data available.
One consistent theme is that the most productive races and populations arose from hybridization of previous races. Older races developed in isolated regions were brought together by the movement of people and natural hybridization occurred. It seems that a certain fraction of heterosis obtained from crossing of races persisted in future generations to produce more productive races (see Chapter 10). The highly productive US Corn Belt Dents is a good recent example of this phenomenon. Crosses of races or populations that are considered promising will require time to allow for adequate recombination. The time required to evolve new races by hybridization will fit long-term objectives and the payoff may not be obvious to the breeder who initiated the program. The ‘hodgepodge’ (Brown, 1953) or ‘mess’ (Wellhausen, 1956, 1965) of hybridization will require mild selection with time to meld the favorable factors of parents in the population derived by hybridization.

Because it is believed that maize originated in the Western Hemisphere, maize grown in other parts of the world is in a sense exotic. Hundreds of years of natural and artificial selection in all parts of the world where maize was grown, however, have developed many distinct local races and varieties. Only in the last 50–60 years, when it was realized that a great amount of variability existed within the species *Z. mays*, were organized efforts made to collect, study, preserve, and use potential genetic resources in all parts of the world (Table 11.1). Consequently, in most parts of the world attempts have been made to introduce foreign materials to local breeding programs. The fate of the foreign introductions varied from little or no success to the greatest possible success in replacing local germplasm.

Until World War II most of the European breeding programs developed varieties and hybrids from use of local germplasms. After World War II extensive use of inbred lines from the USA was made in the production and growing of hybrids. The most productive hybrids of Europe involve crossing US dent lines with European flint lines or lately US dent lines with US dent lines. Improvement programs in the European germplasms continue at a similar pace as in the USA. Almost no maize breeding programs existed in Africa until about 40 years ago. Extensive evaluations of exotic races and varieties have been made in the past 30–40 years because local varieties were on the average very poor. Introduced varieties that showed a much better performance than local types were used effectively for population improvement. Breeding programs using local and exotic germplasm are currently active in most of Africa with guidance and assistance being provided in some instances by the international research centers and greater interest by the private sector. In addition to the drought tolerance concerted efforts on inbred–hybrid systems, population–hybrid programs (Carena, 2005; Carena and Wicks III, 2006) could be cost-efficient solutions for the cooperative systems currently in place.

The majority of races and varieties of maize are found in Central America and South America. Practically all natural genetic variabilities for adaptations to all latitudes and altitudes and germplasms for special and general purposes can be obtained in this area. Thus in this area the interchange of exotic material should result in the best success. Local habits and social conditions present some limitations for use of exotic material. In many regions maize is used for human consumption and needs to fit rigid local standards. For instance, the famous Cuzco material, with
its eight-rowed ears, very large kernels, and very late maturity, is widely used and not easy to replace. People do not easily accept a different product even if it has proved to be more productive or has some other agronomic advantages. In these instances the use of exotic material to include some genetic variability without changing the plant and ear patterns would be a very long process. In Argentina substantial progress has been made where most of the maize is grown as hybrid. There has always been a policy, however, to produce the dark orange flint material to take advantage of the premium prices in the international market. As a consequence very little use has been made of exotic materials, although US dent materials perform quite well and give higher yields than the local flint hybrids. There has been, however, a tendency to move to softer types of kernels (e.g., dents and/or semi-dents) in order to obtain greater productivity. However, the lack of extensive germplasm improvement for local conditions (e.g., lack of public national research infrastructure) has caused negative outcomes against disease outbreaks (e.g., mal de Rio Cuarto). Eyherabide et al. (2006) suggested the need for improving the flint heterotic group. Chile has almost completely adopted US hybrids with great success. Since most of the maize is grown for animal feeding, there was no objection to use of the higher yielding US yellow dent hybrids. During the Civil War many Americans migrated to Brazil bringing samples of US yellow dent maize, which were crossed with the local orange flint Cateto and resulted in many types of dent maize. During 1910–1915, maize shows similar to those in the USA were organized in some areas of Brazil, and new introductions were made. These materials were also used in the natural formation of local dents. The first hybrids developed were of Cateto germplasm (orange flint kernels) and were not highly productive. Semi-dent hybrids were subsequently obtained with a little increase in performance. Experience showed that local flint and dent germplasms were generally poor with low possibilities for inbred line development, especially the dent material. Only after the 1950s were organized introductions performed in Brazil. Tuxpeno, a race from Mexico, substantially increased productivity in population improvement programs and later was used in the development of inbred lines that resulted in superior commercial hybrids. Significant contributions were also achieved by the use of flint germplasms, especially the ones from Colombia (ETO material) and Cuba. Brazil is one of the countries in which all local varieties had poor performance and exotic germplasm gave a substantial improvement. It is estimated that exotic germplasm is responsible for 100% improvement of yield in relation to the local germplasm.

Wellhausen (1978) reported yield data for 10 dent or semi-dent varieties crossed with each of 10 flint or semi-flint varieties. The crosses were tested at six sites in Latin America. From these trials and others, Wellhausen (1978) identified four racial complexes that he considered useful for immediate improvement of maize in the tropics:

1. Tuxpeno and its related Caribbean and US dents
2. Cuban Flint,
3. Coastal Tropical Flint, and
4. ETO.
Tuxpeno is the only pure flint and it was considered to be outstanding because of its yield capacity, exceptional vigor, and resistance to common maize diseases. Tuxpeno should also be useful in US breeding programs because it is generally considered to be one of the putative parents of Corn Belt Dents (Fig. 11.1). Crosses of Tuxpeno with Cuban Flint, Coastal Tropical Flint, and ETO were high yielding and expressed considerable heterosis. Cuban Flint, Coastal Tropical Flint, and ETO are flint complexes that showed good disease resistance and had good yield potential. Wellhausen (1978) stated that the four racial complexes had been used extensively in maize breeding programs in the lowland tropics in the past 20 years.

The greatest diversity of maize germplasm is located in the subtropical and tropical area of Central and South America. This wealth of germplasm probably carries unique alleles not present in the sequence of, for example, B73. Before the tropical germplasm resources can impact the major temperate areas of maize breeding programs, it is necessary to select for adaptability to temperate environments and agronomic traits (particularly root and stalk strength) important for modern large-scale maize production. Distinct methods have been used to introduce, adapt, and improve exotic germplasm for temperate areas:

1. Inbred lines included in tropical hybrids (Goodman, 1985, 1999a).
2. Selection within tropical varieties that are considered important sources of lines for tropical hybrids (Hallauer, 1999a).
3. Adaptation of elite tropical populations and inbred lines derived from elite early lines × elite late GEM breeding crosses to early-maturing regions (Carena et al., 2009a).

In all instances, pre-breeding is necessary to develop either inbred lines or populations that can contribute useful genes for temperate areas. Hallauer and Carena (2009) summarized the pre-breeding activities of exotic germplasm reported by the different maize programs. No coordinated programs for the preservation, introduction, and enhancement of exotic germplasm were in place until the 1980s. Under the leadership of W. L. Brown, the Latin American Maize Program (LAMP) was initiated to rescue, regenerate viable seed, and evaluate germplasm accessions in Latin America. There were 12,113 accessions evaluated and after five stages of evaluation, 268 accessions were identified that were considered elite germplasm sources (Pollak, 2003). After the completion of LAMP, it was quickly realized that additional breeding effort was needed before the identified accessions would be of possible interest and contribute to broadening the germplasm base of temperate area breeding programs. To accomplish the pre-breeding phase of the LAMP program, the US Congress appropriated $500,000 annually to conduct a program designated as Germplasm Enhancement of Maize (GEM). The ultimate goal of GEM is to broaden the genetic base of the hybrids grown by the US producers (Pollak, 2003). The GEM project is a cooperative effort that includes personnel employed by both the public and private sectors of research. Their effort concentrated on central and southern US areas. In order to expand GEM efforts, North Dakota State University (NDSU) initiated a long-term program to increase the
Potential and Use of Exotic Germplasm

Genetic diversity of hybrids in the northern USA with the rapid incorporation of elite exotic germplasm 10 years ago (Carena, 2010). The goals were to move elite tropical and temperate maize germplasm northward. This was achieved by adapting LAMP-GEM germplasm to short-season drought and cold-prone environments (NDSU EarlyGEM program) and developing new and unique early-maturing lines for industry use (Fig. 11.2).

![Diagram of breeding process]

**Fig. 11.2** Long-term breeding process followed to develop NDSU EarlyGEM lines adapted to North Dakota
A modified pedigree selection program was initiated including elite GEM and ND lines to move US Corn Belt GEM germplasm northward and westward. The inspiration of this long-term adaptation program was the backcross breeding program initiated by Dr. Pinnell at the University of Minnesota (Rinke and Sentz, 1961) that yielded successful Minnesota ‘A’ lines (e.g., early versions of B14). However, the difference was that early-maturing parents were the recurrent ones and that only one backcross generation was produced without the screening of $F_2$ segregating populations in order to reduce the time for exotic inbred line development (Carena et al., 2009b). Approximately 5,000 lines from nine BC$_1$ populations were selected for inbred line development and early-generation testing. Stiff Stalk donors (CUBA117:S1520-388-1-B, CHIS775:S1911b-B-B, and AR16026:S17-66-1-B) and non-Stiff Stalk ones (BR52051:N04-70-1, SCR01:N1310-265-1-B-B, FS8B(T):N1802-35-1-B-B, UR13085:N215-11-1-B-B, CH05015:N15-184-1-B-B, and CH05015:N12-123-1-B-B) were advanced through the NDSU pedigree selection process (see Chapter 1) including drought tolerance screening under managed stress environments (Carena et al., 2009b). Foundation Seed Company testers representing the B14, Iodent, and unrelated early-maturing heterotic group testers were utilized for early and late-generation testing. Results are encouraging and first NDSU EarlyGEM lines are scheduled for release in 2010–2011. Results agree with the finding that maize inbred lines with certain percentages of non-Corn Belt germplasm often have combining abilities for grain yield greater than inbred lines carrying 100% of US Corn Belt Dent germplasm (Griffing and Lindstrom, 1954) with the advantage that exotic populations carry more genetic diversity (Goodman, 1965b). Therefore, incorporating exotic alleles seems to have greater chances for significant success in maize than, for instance, certain self-pollinating crops (Carena et al., 2009a). Earlier maturing versions of GEM lines (12–20 days earlier) were produced. Recovered BC$_1$:S$_1$ lines could retain grain yield at ND relative maturity in new hybrid combinations. Preliminary data showed several experimental NDSU EarlyGEM lines (theoretically 12.5% exotic) have, among other traits, similar and/or better grain yield performance as testcrosses than popular industry hybrids at similar grain moisture at harvest. Rinke and Sentz (1961) considered their work as one of the most outstanding achievements in 20 years of breeding work and both research programs seem to be a good utilization of germplasm resources as a consequence of public–private cooperation. This is the first research devoted to <90RM germplasm enhancement with the incorporation of tropical and late-temperate GEM genetic materials for unique inbred line development. The most advanced earlyGEM lines available have shown above average drought and cold tolerance, grain quality, lodging resistance, fast dry down, and tolerance to ear rots (especially after 2009 season) (Carena et al., 2009a,b). Winter nurseries with three generations per year are providing new lines in 4 years. In addition, new early-maturing earlyGEM breeding populations are being developed to increase efforts to identify alternative heterotic patterns for the northern US Corn Belt and continue the development of new and unique lines currently not present in the industry. Long-term cooperation within GEM will continue to provide new germplasm for the northern USA.
Because of possible global climate changes, drought-tolerance may become a major challenge for maize production. Western North Dakota is not the exception. In fact ethanol plants have initially established where relatively cheap energy sources are present in the west. The development of elite maize hybrids with improved stress tolerance is essential to obtain maximum yields and minimum costs associated with production and the environment. Drought tolerance is genetically a very complex trait with large genotype by environment interactions. Therefore, marker-assisted selection and transgenic approaches are challenging and might be limited solutions to the problem. Non-transgenic breeding approaches have a potential of significantly increasing genetic progress without limits and exploiting polygenic effects compared to transgenic approaches exploiting single-gene effects. Data of early- and late-generation hybrid trials in western ND and eastern Montana (MT) showed that hybrids including ND experimental lines, developed under drought-managed winter nursery environments, had significantly better yield, lower grain moisture at harvest, higher test weight, higher extractable starch, higher fermentable starch, higher oil, and higher protein than industry check hybrids (Carena et al., 2009b). At least 40 NDSU experimental hybrids yielded better than checks in western ND with specific hybrids exceeding 193.8% improvement. Several of these have exotic alleles in their genetic background as they belong to our long-term NDSU EarlyGEM tropical and temperate inbred line development program for short-season environments. There is no limit to genetic improvement for drought tolerance when most tolerance genes are targeted in the breeding process. Heterotic effects are unique for each hybrid and sequencing efforts on only B73 may limit the identification of useful alleles for drought tolerance and other complex traits. One of the reasons behind the significant current genetic improvements is the lack of early-maturing drought-tolerant hybrid checks. The NDSU maize breeding program has the desire to use proposed transgenic drought-tolerant industry hybrids as checks for western ND and eastern MT short-season environments when they finally become available but this is an area not served by the industry yet and the public sector is still essential for areas where the market is not large enough for significant industry investment.

Germplasm enhancement is timeless: i.e., consistent, incremental genetic improvements require long-term programs to ensure consistent genetic improvement. The GEM project seems to be a necessary program for all interests of maize improvement. GEM should provide the necessary mechanisms to provide useful exotic germplasm for broadening the genetic base of temperate breeding programs, if funding continues to support the research.

It is generally acknowledged that successful breeding programs have either developed or had access to elite germplasm. Sources of elite germplasm, therefore, are imperatives for successful development of inbred lines and hybrids. Although the availability of elite germplasm is recognized, very limited, concerted efforts, however, have been given to the evaluation and improvement of germplasm resources. During the past 60 years, only a limited number of researchers and agencies or organizations have allotted significant resources to the evaluation and improvement of germplasm for breeding purposes. Harlan (1975), Kahn (1985), Raeburn (1995), and Evans (1998) have discussed the evolution of our important
cultivated crop species from wild weedy plants; their importance in our modern cultivated plants; and the importance for the collection, maintenance, evaluation, and improvement of our germplasm resources. But germplasm has received only minor attention in comparison with development of inbred lines for testing in hybrids. As reviewed in this chapter, there have been numerous reports of germplasm enhancement, but most have either been discontinued or downsized with the retirement of individuals or changes of research emphasis by the organizations and/or agencies. The more-or-less haphazard emphasis over time has not contributed to the consistent genetic improvement of our germplasm resources. It is encouraging that the LAMP and GEM programs in the Western Hemisphere address the concerns of maize germplasm for breeding purposes. To ensure that past efforts are not negated, it is very important that GEM, and similar programs, be properly funded and managed to ensure future genetic improvements of products provided to the producers. In addition, private-public partnerships to share breeding rights access and expand maize production in challenging niche environments will not only improve industry germplasm but also will secure dependable hybrids to farmers.

References

Brown, W. L. 1960. Races of Maize in the West Indies. NAS-NRC Publ. 792, Washington, DC.
References


Paterniani, E., and M. M. Goodman. 1977. Races of maize in Brazil and adjacent areas. CIMMYT, Mexico.


References


Wallace, H. A. 1923. Burnett’s Iodent. *Wallaces’ Farmer* 46(6), February 9, 1923, Des Moines, IA.


Adaptation of germplasm and maximization of its germplasm improvement are necessary pre-breeding goals for any applied maize breeding program developing cultivars. Recurrent selection methods can contribute to meeting the goals of continuous genetic improvement for significant genetic gains. These methods are complementary to inbred line development methods (e.g., pedigree selection, doubled haploids) as they provide improved progenies cyclically, the same way half-sib progenies provided B73. They will not replace other breeding methods, but should be integrated with them (Hallauer 1981, 1985, 1992; Pandey and Gardner 1992; Carena and Wicks III 2006; Carena, 2008; Hallauer and Carena, 2009). However, very few breeding programs are still active with intra- and inter-population recurrent selection programs.

Cultivated maize arose from its wild ancestors as a result of mutation, natural and artificial selection, and adaptation to the environment. Plant breeding is the art and science of improvement of crop plants to meet the needs of people. Processes of evolution are involved, but they are directed by people to hasten the attainment of their projected goals. Plant breeding is a broad discipline that requires some level of competence in genetics, botany, statistics, pathology, entomology, and an appreciation of the forces of the environment on plant growth and development. The relative importance of the art and science of plant breeding is ill-defined, but plant breeding has played an important role in the development of crop plants since the hunting and gathering stage of humans because of their dependence on plants for survival. Although people have always attempted to direct the evolution of plants to meet their needs, not until the genetic laws of Mendel and the principles of randomization and replication were understood and developed did plant breeding as a science become prominent (Carena and Wicks III, 2006; Hallauer, 2007; Hallauer and Carena, 2009).

The principles of maize breeding as we understand them today have been developed during the 20th century, beginning with the publications of Shull (1908, 1909, 1910) and East (1908). The change from use of open-pollinated varieties to use of double-cross hybrids was a significant advancement in developing maize plants with improved standability and greater yields. The impact of inbreeding and hybridization techniques on yields is shown in Fig. 12.1.
Fig. 12.1 National average maize yields in the United States from 1965 to 2008 (Courtesy of Dr. A.F. Troyer)

The national average for US maize yields was consistently low until the latter part of the 1930s. Yields did not fluctuate greatly and average US yields attained 18 q/ha in only 2 years before 1935. Small but consistent improvements have been made since the introduction of double-cross hybrids in the 1930s to meet the changes and challenges of the improvements of maize husbandry and mechanization. Yield increases shown in Fig. 12.1 reflect the rapid acceptance in the use of hybrid seed, changes in types of hybrids grown, improvements made in crop husbandry, genetic improvements of parental inbred lines by recycling, and the rapid changes made in molecular genetics to develop tolerance to pests and weeds (Fig. 12.1). GMO maize was, in all instances, dependent on improved genetics by traditional breeding methods on unique germplasm.

Evidence that genetic improvements were made in maize hybrids was reported by Duvick (1977, 2004) and Russell (1974, 1986) in the US Corn Belt. Both authors included hybrids produced and grown in the decades since the introduction of hybrids in replicated yield trials. Each report provides evidence of significant genetic improvement in yield. Duvick included two sets of hybrids in his trials:

1. Hybrids grown in the five decades, including double crosses grown in the 1930s, 1940s, and 1950s and single crosses grown in the 1960s and 1970s.
2. Single crosses produced from inbred lines included in the hybrids for the five decades.

Russell included hybrids (double and single) that were used for the five decades. Both studies included three plant densities, which were intended to simulate plant
population densities used for growing early hybrids, plant densities presently used, and plant densities greater than generally used in the US Corn Belt at the present. All comparisons showed that genetic improvement had been made in the development of hybrids to meet the conditions under which they were grown. Russell calculated that gain in hybrid performance attributable to breeding was 63.2% and Duvick determined that proportion of total gain due to breeding was 57 and 60% for his two sets of hybrids. Gains were in good agreement for the three sets of experiments. Russell also found that if we were still growing maize at plant population densities commonly used for the first double crosses, we would realize little yield improvement over the first experiment station double crosses. Hence genetic improvement depended also on planting practices and the tolerance of these hybrids to higher planting densities.

Response of hybrids for yield improvement depended on correlated responses for improved roots and stalks, improved disease and insect resistance, and resistance to barrenness at higher plant population densities across decades. Modern maize production requires hybrids that have acceptable standards of root and stalks and ear attachment to permit retrieval by mechanical harvesters. Grain yield itself is the result of the total genotype of the hybrid in modern maize growing conditions. Hence, harvestable yield is the standard of measure of a hybrid rather than the potential genotypic yield which can be measured by hand harvesting to collect all ears whether they are on standing or broken plants. If there are indeed yield genes, they must be incorporated within genotypes that include genes contributing to good roots and stalks, strong ear attachment, acceptable maturity, and good health and vigor. Additional requirements for modern hybrids reduce selection intensity for the specific trait (e.g., yield). Therefore, inclusion of additional traits broadens our definition of yield. In spite of restrictions imposed for developing superior yielding hybrids, significant progress has been made (Fig. 12.2). Plant breeding is dictated by economic uses of crop species and selection techniques have to be adapted to meet demands.

Fig. 12.2 Predicted yields of hybrids at 10-year intervals based on actual yields of the two best hybrids in each double-cross group (Russell, 1974)
Although it seems slow, continuous genetic improvements have been made in developing hybrids. The concern is whether the same progress can be achieved in the future. Progress in the development of mechanized equipment has improved the timeliness of planting and harvesting; extensive applications of fertilizer (particularly nitrogen) have been common since 1950; plant population densities have been slowly increasing; development and use of pesticides have increased dramatically to provide a better environment for maize growth and development; and management abilities of farm operators have developed to a high level of sophistication including global positioning systems. All these developments and refinements in concert with genetic improvement have contributed to higher yields. Many of the factors, however, may have attained their plateaus so that dramatic changes may not be forthcoming, but evidence to present does not suggest that yield plateaus have been attained. There have been suggestions that the potential of molecular genetics in maize breeding will double maize yield by 2030, or 18–19 Mg/ha. The sequence of the B73 genome has been completed and more genotype sequences need to target unique alleles and representative samples. They could permit a better understanding of the genetic basis of heterosis and/or the genetic traits contributing to greater drought tolerance. In each instance, transgenic events can be identified and incorporated for the traits of interest. Because of the concerns of environmental quality, possible climatic changes, water limitations, and finite land resources, maize breeders and producers may be confronted with more restrictions in selection that emphasizes greater grain yields. For instance, less fertilizer will be used in the future than in the past or at the present and breeders may develop hybrids tolerant to low nitrogen requirements due to its elevated cost. Development and use of pesticides may be restricted because of the concerns of environmental quality. Potential limitations on further cultural and management changes and improvements emphasize that continued genetic improvements are imperative. However, research on interactions is encouraged before duplicating breeding programs (Carena et al., 2009c) and/or evaluating only at certain environments (e.g., organic) or only certain genotypes (e.g., evaluating genotypes bred under high N levels to identify genotypes with low N needs). Exploiting unique and challenging environments can accelerate adaptation to climate change.

It seems that genetic improvement of maize can be maintained by expansion and refinement of breeding techniques, integration of cyclical selection programs with applied breeding programs, and greater usage of germplasm available. The information summarized in Chapters 5, 7, 8, 9, 10, and 11 represents approximately a century of research conducted since the suggestion of the pure-line hybrid concept by Shull (1908, 1909, 1910). The precepts of Shull with some minor modifications have been essentially used in most maize breeding programs. Population improvement schemes and maize germplasm available have not been used extensively. Dudley and Moll (1969) and Moll and Stuber (1974) have discussed interpretation and uses of genetic variances for populations and relevance of results. Integration of these facets with those currently used is required if we are to continue to have genetic improvement. None will result in any spectacular leaps forward, but integration of all phases will ensure a slow, steady rate of genetic improvement (Hallauer 1985, 1992).
Maize applied breeding programs include three important phases to meet the short-, intermediate-, and long-term objectives:

(1) Choice of germplasm.
(2) Cyclical improvement of germplasm chosen.
(3) Development of lines for use as parent stocks in production of single-cross hybrids (for areas requiring hybrids) and development of improved varieties, synthetics, and composites and their own hybrids (for areas where hybrids are not currently practical) (Dowswell et al., 1996; Carena, 2005; Carena and Wicks III, 2006).

All phases are equally important, but resources and time expended for each may vary considerably.

12.1 Choice of Germplasm

Germplasm to include in the breeding program may involve

(1) choosing from current sources of germplasm and information available, or
(2) collection, development, and evaluation of germplasm before choices are made.

In the first instance the breeder depends on the information and opinions of others. This approach may be adequate if the evaluations were made in an environment that is similar to that in which the germplasm is to be used. If the information was obtained from a very different environment, the germplasm may not be adapted to the environment desired.

The collection, development, and evaluation of germplasm require several seasons, but the breeder has the opportunity to directly observe the response of the germplasm to the particular environment. Several collections may be obtained from germplasm banks and are grown to determine which ones have the best plant and ear development and adaptability. Similar collections may be combined to form a composite or a few may be sufficiently distinct from one another and may have sufficient genetic variation for direct use. Either approach may be used, but generally a combination is used. Some evaluation trials may be conducted and used in conjunction with previously available information that in many instances may be experiences of other breeders.

Choice of germplasm is a critical decision that requires considerable thought. Hasty decisions either to eliminate or to decrease the number of growing seasons required may in the long run increase the number of growing seasons required to develop usable materials. Germplasm chosen forms the basic material of the breeding program. Germplasm that has a low frequency of alleles for the traits desired may require either several additional growing seasons or greater samplings to develop desirable genotypes or populations. Unfortunately, information available
to the breeder for the large collections of germplasm is limited which emphasizes the importance of considering all the information available before choosing. In many instances the selected germplasm will be the basis of the breeding program for the lifetime of the breeder or it could be the basis for genetic experiments and increased knowledge (e.g. NAM population). Therefore, accurate sampling and germplasm choice before embarking in maize genetics and breeding experiments is key. Choice of germplasm will determine maximum potential improvement that can be attained via breeding while the breeding system used will determine how much of that maximum potential can be realized. In breeding programs with extensive resources, large volumes of molecular markers are generated among inbred lines to assist breeders in choosing parents that seem to have greatest potential for selection within F2 populations to develop superior progenies.

12.2 Recurrent Selection and Germplasm Improvement

After choice of germplasm, the next procedure is to use some type of cyclical selection program to maximize the genetic improvement of desired traits. The breeder can proceed immediately to extract materials that are planned for use by the farmer; this is a natural, practical response. The important aspect is that superior materials selected for the breeding nursery should also be recombinated to re-synthesize the basic population. If considerable thought, time, and expense have been devoted to choice of germplasm, the logical process is to continue using the germplasm in the breeding program. Presumably the superior progenies isolated from the basic population included the genetic factors that meet standards imposed by the breeder. Probably none will meet the desired standards of all traits, but at least the selected progenies are better than a random sample. Recombination of superior progenies will increase frequency of favorable alleles, which in future cycles will increase the opportunities of isolating a greater frequency of progenies that meet desired standards for most traits. Similar methods apply to the recycling of elite inbred lines in breeding programs dedicated to developing inbred lines and hybrids (Hallauer, 1985). A simple example of the possibilities of recycling is illustrated in Fig. 12.3. In this idealized example for finite populations, expected variability of single crosses of original and improved populations is the same but the mean has increased, and the best possible hybrid extracted from the improved population is superior to that of the original population. The example in Fig. 12.3 usually is not realized after only one recycling. If the recycled population represented 10 cycles of selection and

![Fig. 12.3 Expected distribution of single crosses from the original and improved populations (Eberhart, 1970)](image)
12.2 Recurrent Selection and Germplasm Improvement

Recurrent selection and germplasm improvement are essential processes in plant breeding that can significantly enhance crop performance. Recombination, the improvement for yield would be striking. Instead of intermat ing superior progenies, the breeder may be tempted to resample the original population. If sampling was adequate, re-sampling would be expected to give progenies that have the same relative distributions. Changes in relative distribution and mean of progenies extracted after only one to three cycles of selection may not be detectable. It is essential that recycling be continued on a regular basis over a period of time.

Results of recurrent selection programs conducted in maize are summarized in Chapter 7. Effectiveness of recurrent selection depends on genetic variability and allele frequencies within original populations and heritability estimates of traits under selection. For example, several germplasm improvement programs for pest resistance were conducted before transgenic maize (e.g., Bt) was available. Penny et al. (1967) reported that three cycles of inbred progeny recurrent selection for first brood European corn borer resistance were effective in developing populations that had an acceptable level of resistance (Fig. 12.4). Based on a rating scale of 1 (resistant) to 9 (susceptible), the mean of 484 S₁ progenies for the C₀ population of five varieties was a high intermediate (5.5), whereas the mean of the 484 S₁ progenies for the five C₃ populations was within the resistant range (2.5). Distributions of S₁ progenies from the original and last cycles of selection also showed that three cycles of recurrent selection seemed adequate for establishing a greater level of resistance. Frequency of resistant S₁ progenies for the C₃ cycle was skewed to the resistance classes with none in the 7, 8, and 9 susceptible classes.

Similar recurrent selection procedures were used by Jinahyon and Russell (1969a) for developing resistance to Diplodia zeae in the open-pollinated variety Lancaster Surecrop, which has a high level of susceptibility to stalk rot organisms. One of the methods of evaluating response to selection was use of 100 S₁ lines from the C₀ and C₃ populations. Distributions and means of these lines are shown in Fig. 12.5.

---

**Fig. 12.4** Distributions of S₁ progenies for the original population (C₀) and after three cycles of recurrent selection (C₃) for first brood European corn borer resistance (Penny et al., 1967)
Fig. 12.5  Distributions of S₁ progenies for the original population (C₀) and after three cycles of recurrent selection (C₃) for *Diplodia zea* resistance (Jinahyon and Russell, 1969a)

Ratings were made on a scale of 0.5 (resistant) to 5.0 (susceptible). Average stalk rot ratings of the 100 S₁ lines were 4.1 for the C₀ population and 2.4 for the C₃ population. An important feature of recurrent selection for stalk rot resistance was the change in distribution of S₁ progenies. Very little overlap occurred between the two distributions; the C₃ population had very few S₁ progenies exceeding a rating of 3.0 and very few C₀ S₁ progenies had a rating less than 3.0.

In both instances, three cycles of recurrent selection based on S₁ progeny evaluation were effective in changing the mean and distribution of the populations for the trait emphasized in selection. Heritability of first brood European corn borer and stalk rot resistance is greater than heritability for yield because techniques were developed for artificial infestation and infection of these two important pests. Development of artificial means for minimizing the frequency of escapes improves the heritability and, consequently, the effectiveness of selection. Response to selection would be expected to be greater than for traits that depend on natural infestation or infection and the effects of the environment on the establishment and development of the pest. If the techniques are available, it seems that recurrent selection is an effective method for improving mean performance of a population and chances of obtaining superior progenies from improved populations. Results for first brood European corn borer and *Diplodia zea* stalk rot agree with the hypothetical example illustrated in Fig. 12.3. Three or four cycles of recurrent selection seem sufficient in most instances in developing populations that have acceptable levels of resistance (Hallauer, 1973a). Although Penny et al. (1967) and Jinahyon and Russell (1969) were successful in selection for greater levels of resistance to two important maize pests, the emphasis of selection on only pest resistant resulted in decreased grain yield and usually undesirable changes in other agronomic traits (Devey and Russell,
Recurrent Selection and Germplasm Improvement

Selection indices for multi-trait, multi-stage, and multi-environment selection would be a solution to this problem (Hallauer and Carena, 2009).

In contrast, Penny and Eberhart (1971) reported on 20 years of reciprocal recurrent selection (RRS) for yield in two synthetic populations, BSSS and BSCB1. Twenty years were required to complete four cycles of recurrent selection for yield whereas only 4–8 years would be required for first brood European corn borer and stalk rot resistance. Four cycles of RRS improved the population cross (1.18 ± 0.24 q/ha) and BSSS (1.38 ± 0.62 q/ha), but no improvement was realized in BSCB1 (−0.64 ± 0.62 q/ha). Although results were not striking and may seem discouraging, they were not totally unexpected because of limited opportunities (four in this instance) for selection and recombination. The number of selection cycles could have been increased if off-season nurseries were available. Russell and Eberhart (1975) tested the population cross as well as the crosses among five selections from BSSS and BSCB1 after five cycles of RRS. The five selections from BSSS and BSCB1 were in the S2 generation and were included among the 10 selections used to synthesize the sixth cycle in each population. Average yield of S2 line crosses exceeded the population cross in all instances and the best S2 line cross exceeded the population cross by 35%. Two of the S2 line crosses yielded significantly greater than B37 × Oh43, a single cross that had been used in commercially grown hybrids. Keeratinijakal and Lamkey (1993) evaluated response to selection for grain yield after 11 cycles of RRS for BSSS and BSCB1 populations and direct response was 7.0% per cycle of selection and level of heterosis increased from 25.4% for the original population cross to 76.0% after 11 cycles of selection.

Moll et al. (1977) reported that six cycles of RRS improved chances of obtaining superior single crosses from the two source populations, Jarvis and Indian Chief. They examined frequency distributions of single crosses derived from unselected lines developed from original and improved populations, which had been improved by RRS (Fig. 12.6).

Average yield of single crosses from the improved populations was 12.5% greater than those from the original populations. RRS improved the means of the hybrids but the distributions of the crosses were similar, which indicates that chances of obtaining a superior hybrid are greater in the improved populations than in the original populations. The 10 best single crosses from the selected populations also averaged 8.6% greater yield than the 10 best single crosses from the unimproved populations.

Hallauer (1984) and Eyherabide and Hallauer (1991) have reported response to reciprocal full-sib selection to develop genetically improved source populations for developing lines and hybrids. The frequency distributions for the C0, C7, and C13 cycles of selection show the same trend (Moll et al., 1977).

Relative to the mean of the six check hybrids, the distribution of the C0 and C13 cycles was transposed. Only two of the 144 C0 full-sib families exceeded mean of check hybrids, whereas only 17 of the 214 C13 full-sib families had yields less than the mean of check hybrids. Eyherabide and Hallauer (1991) reported 7.5% gain per cycle for grain yield after eight cycles of RRS for BS10 and BS11 and heterosis increased from 2.5% for the C0 × C0 population cross to 39.6% after eight cycles of selection.
Available evidence suggests that recurrent selection procedures increase opportunities for obtaining superior progenies and crosses. In most instances only a limited number of cycles of recurrent selection have been completed, but the results are encouraging. Reported results tend to conform to the simplistic model illustrated in Fig. 12.3. The primary objective of recurrent selection procedures is to gradually increase frequency of desirable alleles for breeders’ primary objectives while maintaining genetic variability. Constant selection pressure and recombination of progenies that possess genes meeting the standards of desirability established by breeders will develop improved breeding populations that will enhance chances of isolating superior lines and hybrids as happened with B73 and their derived hybrids. If the maize program has not developed to the point of use of lines and hybrids, continuous recurrent selection procedures can develop cultivars (e.g., population hybrids) to be used directly by farmers (Dowswell et al., 1996; Carena, 2005; Carena and Wicks III, 2006). The relative success of recurrent selection depends on the complexity of the trait under selection, experimental techniques available for screening progenies, and effects of the environment. Comparisons of results show that significant improvements can be expected for most traits but the rate of improvement is greater for some traits than for others.

Yield is usually the most important economic trait considered in maize breeding programs but for certain environments (e.g., North Dakota) other traits (e.g., grain moisture at harvest, fast dry down, cold and drought tolerance) might be as important as yield. Use of recurrent selection techniques for yield improvement has not been as impressive and consistent as for traits that have a greater heritability and for
which techniques are available to minimize the effects of the environment. Yield is a specific measurable trait that is a composite of the plant genotype in response to the environment. Good vigor and health in a high-yield environment result in a high yield, but this situation is usually the exception. Good vigor and health, or freedom from maize pests, usually have greater heritability estimates than yield. It seems that our breeding populations can be improved for these traits to the point that economic losses are not serious in a relatively short time, perhaps 4–8 years. Then emphasis can be given to selection for improved yield. This suggests a two-stage selection program, but as described later, the selection of vigor and health traits can be combined with recurrent selection for yield improvement. Combining selection of as many traits as possible into one program will save time because the two-stage, or tandem, selection increases the time required. Recently, molecular biologists have enhanced resistance to European corn borer and corn rootworms and tolerance to herbicide applications by the insertion of elements into elite lines used to produce hybrids. Current genetic improvement efforts could be complemented with marker-assisted selection on traits that are difficult to measure and have low heritability. Any applied maize breeder would be happy to utilize, for instance, accurate markers for root traits as often selection for seedling vigor is not accompanied by selection for better root systems that are below ground. Fast dry down could be another potential trait although a new method of measurement for inbreds and hybrids has been proposed (Yang et al., 2010) and new doubled haploid lines are being developed in cooperation with industry. Unfortunately, marker efforts on these types of traits have not been the priority among research groups.

Success from multi-trait selection depends on level of genetic correlation among traits included for selection (see Chapter 7). If genetic correlations are low, simultaneous selection for different traits will cause no difficulties. Low genetic correlations are an advantage to the breeder who wants to combine good vigor and health traits with improved yield. Strong positive correlations between desirable traits are of greater advantage. Experience and data seem to indicate that many of the vigor and health traits do not have a strong genetic correlation with grain yield. Jinahyon and Russell (1969b), for example, found that the selection imposed for improved stalk rot resistance (Fig. 12.5) did not reduce yield, but continued selection for four additional cycles did impact grain yield negatively (Devey and Russell, 1983). The associated changes with earlier cycles of selection were greater plant vigor, later maturity, better disease resistance, greater stalk strength, and greater yields in hybrids. Selection for greater yield and vigor in temperate regions tends to develop materials having later maturity. Unless one emphasizes selection for earlier maturity with selection for greater yields, the populations tend to become later with associated changes of higher ear placement and taller plants.

Hallauer (1978) and Eyherabide and Hallauer (1991), from use of reciprocal full-sib selection for harvestable yield in BS10 and BS11, found that grain yields were significantly increased in the populations themselves and in their crosses after three and eight cycles of selection for yield. Associated changes showed that stalk lodging was significantly reduced, grain moisture was reduced, and no change occurred in root lodging. Selection pressure was also effective for other traits in desired
Breeding Plans

directions. In reciprocal full-sib selection, however, selection was based on selfed progenies and full-sib crosses. Carena and Hallauer (2001a), Carena and Cross (2003), and Hyrkas and Carena (2005) found similar results on multi-trait intra-population recurrent selection programs on Leaming and NDSAB, respectively (Figs. 12.7 and 12.8). Also, stratified mass selection for adaptation has produced significant and desirable correlated responses even though selection has been for one trait (Hallauer and Carena, 2009).

![Graph 1](image1)

**Fig. 12.7** Indirect responses of grain yield and of stalk lodging in Leaming open-pollinated maize variety after three cycles of $S_1$–$S_2$ recurrent selection (Carena and Hallauer, 2001a)

![Graph 2](image2)

**Fig. 12.8** Direct responses of grain yield and stalk lodging in NDSAB maize synthetic variety after 10 cycles of modified ear to row selection (Hyrkas and Carena, 2005)
Selection pressure for only one trait, such as those traits reported by Penny et al. (1967), Jinahyon and Russell (1969a), and Hallauer and Carena (2009), probably will be more effective than in combination with other traits if we consider the time interval for the one trait. Three cycles of recurrent selection for first brood European corn borer and stalk rot resistance developed populations having acceptable levels of resistance. If we combine selection for these traits with selection for yield, some trade-offs usually are made in the final selections for recombination. As shown by Devey and Russell (1983), Klenke et al. (1986), Dudley and Lambert (2004), and Hallauer et al. (2004) selection that emphasizes only one trait is effective but it is usually detrimental to the overall agronomic performance of the germplasm under selection. Instead of three cycles, it may require three to six cycles of selection to attain a comparable level of resistance. Progress will be made, but at a slower rate because of the compromises made in the selection process. On the other hand, selection for adaptation (e.g., stratified mass selection) can emphasize intense selection for flowering in early cycles and emphasize multi-trait selection with yield after adaptation.

Selection for plant vigor and health traits often can be made before selection for yield. In some types of recurrent selection schemes it is convenient to include selection for other traits before conducting expensive yield trials. Selection occurs at different stages of plant and progeny development and undesirable plants and progenies can be discarded before they are considered for yield evaluation. Strong selection pressure can be imposed. However, if the genetic correlations between traits are low, effective selection can be made in conjunction with yield. For instance, Sezegen and Carena (2009) conducted inbred and full-sib progeny selection for cold tolerance at a rate of 1 year per cycle. Progenies were produced in winter nurseries while evaluation and recombination was conducted in summer trials and nurseries. Since yield was not included in the index, emergence percentage and seedling vigor could be screened before flowering across northern US locations. All progenies were grown in the breeding nursery and a procedure named ‘intra-diallel’ was utilized for recombination of top progenies the same season evaluation across locations was done.

Because most breeding programs are concerned with simultaneous improvement of several traits, there was an initial interest in use of the selection index proposed by Smith (1936). The selection index has been shown to be the most efficient method for maximum aggregate genetic progress, provided that

1. reliable estimates of genotypic and phenotypic variances and covariances are available and
2. appropriate economic weights of each trait can be determined.

Williams (1962) suggested the base index in which the traits are weighted only by their economic values. These values, however, are not easily determined for many traits. Pesek and Baker (1969) recognized the difficulty of assigning relative economic weights and proposed a modified selection index using desired gains of the traits. Suwantaradon et al. (1975) compared the use of these three selection index
methods (conventional, base, and modified) in recurrent selection programs for the simultaneous improvement of several traits. From their comparisons of S₁ testing, they recommended use of the modified selection index when relative economic values of traits are difficult to determine. The main prerequisites for use of a selection index in improvement programs are reliability and simplicity. Subandi et al. (1973) examined gains expected from selection based on five selection indices that included three agronomic traits (grain yield, stalk lodging, and dropped ears). The objective of their study was to develop a simple and useful selection index to determine machine harvestable yield. They found that use of selection indices was more efficient in increasing machine harvestable yield than selection based only on yield. Simultaneous selection for several traits is necessary if recurrent selection methods are used. Index selection must be used and applied for increasing efficiency of selection. Smith et al. (1981), for example, developed a selection index that was based on the means and heritability estimates obtained from the analyses of variance for the half-sib, full-sib, and inbred progenies evaluated in different environments. The means for each trait were weighted by the heritability estimates for each trait. The analyses assumed the correlations between traits that were either zero or very low. The final worth of an entry was determined by summation over the traits measured. Baker (1986) discussed the theory and advantages and disadvantages of the different indices proposed, and Hallauer and Carena (2009) also discussed and presented an example of three indices for a half-sib selection trial. In most systems of recurrent selection, selection for other agronomic traits can be made to enhance the future usefulness of improved breeding populations.

Methods of recurrent selection include selection either within one or more populations (intra-population) or between two populations (inter-population). The choice of which to use depends on traits under selection, objectives of selection, and capabilities of the breeding program. Intra-population selection is more appropriate in some instances, e.g., selection to improve stalk quality or adaptability of a particular population. If the primary objective is to improve the population cross, inter-population selection would be more appropriate. Intra-population selection would emphasize selection for additive genetic effects, whereas inter-population selection would enhance selection for non-additive as well as additive genetic effects. Data by Moll and Stuber (1971), Moll et al. (1977), Eyherabide and Hallauer (1991), Keeratinijakal and Lamkey (1993), Carena and Hallauer (2001a), and Carena (2005) show that inter-population selection increases heterosis of the population crosses without changing the heterosis of crosses of populations subjected to intra-population selection. Intra-population selection can be used to improve two populations that express heterosis in their cross for some specific traits (e.g., stalk quality) without changing the heterosis expressed for yield, assuming no correlated responses. Some intra-population methods of improvement are simpler to conduct than inter-population methods and advantages and disadvantages must be considered relative to the primary objectives. Except for reciprocal full-sib selection, progress depends on the increase in frequency of alleles with favorable additive effects for both intra- and inter-population recurrent selection.
12.3 Integrating Recurrent Selection with Cultivar Development

The third important phase of increasing effectiveness and efficiency of maize breeding is to integrate recurrent selection with inbred line development programs. There has been a tendency in the past to consider recurrent selection schemes as basic research studies that have little or no relevance in applied breeding programs, the same way breeders feel toward molecular basic studies today. However, improved populations by recurrent selection are more elite sources of inbred lines and are often released by public breeding programs as such.

The first recurrent selection programs were initiated to obtain information on response of populations to different selection schemes and to determine how the observed response was related to possible types of gene action. Questions were asked concerning the relative importance of overdominance in heterosis, adequate genetic variation in populations for effective selection, appropriate choice of testers to use for selection, and relative importance of general and specific combining ability. Answers were needed to assist in development and use of effective breeding procedures. Summaries of these studies suggested that adequate additive genetic variance was present in maize populations to expect progress from selection (see Chapter 5) and that response from selection was realized in most instances regardless of selection scheme used and trait under selection (see Chapter 7). Summaries by Sprague and Eberhart (1977), Hallauer et al. (1988), and Hallauer and Carena (2009) also showed that response to selection for yield improvement was similar for the different intra- and inter-population recurrent selection schemes.

An important feature of all cyclical selection schemes is recombination of superior progenies. Progenies selected for recombination are determined from either evaluation of progenies themselves or testcrosses in which either remnant seed is recombined or (more commonly) the selfed seed of the plants used in the testcrosses is recombined. The basis for selection is, in all instances except for mass selection, yield of the progenies themselves or their testcrosses. Early-generation testing is conducted in all instances. Because testing is limited in early generations, the relative ranking is not exact but at least the poorest progenies are not consistently included (Rodriguez and Hallauer, 1991). Because progress from selection is usually realized, limitations of early testing in a few environments are overcome to some extent by recombination and additional testing and selection in subsequent cycles of selection. Most evidence indicates that frequency of superior progenies is greater in advanced cycles of selection, which is encouraging to the applied maize breeder.

Each recurrent selection scheme is discussed in reference to how it can be used in conjunction with inbred line and hybrid development programs. Recurrent selection should not be considered as a separate phase of applied breeding programs. Only when recurrent selection is considered an integral part of applied breeding programs will its real benefits be realized. Additional information can be obtained from recurrent selection programs and they should be considered as one of the breeding tools available to maximize genetic improvement. If early-generation yield evaluations are considered a valid basis of selecting superior progenies for recombination to
form the next cycle of selection, superior progenies also should be submitted to pedigree methods of inbred line development. The number of progenies extracted from recurrent selection programs will not be large for each selection cycle, but they will not be random selections and can be elite sources of genetic diversity. Additional progenies that were borderline cases and not included for recombination may be included. In subsequent cycles material will be continuously and regularly generated for the breeding nursery for additional selection and testing. Greatest success will be obtained when the frequency of favorable alleles is fairly high and the frequency of desirable genotypes will be discouraging in the early cycles of selection.

To be most useful, recurrent selection must be an integral part of the breeding program and not a minor part that is conducted only when it is convenient in certain seasons. All breeders have to plan their programs because of the seasonal nature of plant breeding. Recurrent selection has to be included as one phase of the breeding program that is conducted on a continuous and regular basis. Recurrent selection should be considered as applied breeding, not primarily as a basic research program to determine how realized response compares with predicted response. The potential benefits of cyclical long-term selection programs can be rewarding for inbred line development (Mikel, 2006; Mikel and Dudley, 2006). Even though the odds of developing cultivars from recurrent selection programs (and inbred line development programs) is small it only requires one to make impact (Hallauer and Carena, 2009) showing the importance of inbred lines developed from different cycles of recurrent selection in Iowa Stiff Stalk Synthetic (e.g., B73), which has been used either directly in proprietary hybrids or modified by pedigree selection to develop improved strains.

12.4 Intra-Population Genetic Improvement

12.4.1 Mass Selection

Mass selection is not as amenable as other recurrent selection schemes for generating progenies in each cycle of selection. Mass selection is useful for traits that have a relatively high heritability (e.g., maturity or flowering time, adaptation, prolificacy). Because selection is based on individual plants that were sib-mated by open pollination, the precision is not as great as for progenies evaluated in replicated trials. Controlled pollinations can be made, but this would reduce the advantages of mass selection as a relatively simple and inexpensive method of recurrent selection. Harris et al. (1972) have shown that the frequency of superior testcrosses was greater in an advanced generation of mass selection than in the original, unselected population. But it required the development of progenies by controlled pollinations to evaluate mass selection. Ordinarily, half-sib seed of each ear in mass selection is not grown in the breeding nursery because no inbreeding has occurred, and the large numbers (e.g., 20–25,000 plants are used in adaptation to North Dakota conditions) from each cycle may preclude it. Efficiency of mass selection is improved if tassels of
undesirable plants are removed before pollination. Parental control is doubled (see Chapter 6), which increases the efficiency of mass selection. Gardner (1977, 1978) summarized the response to mass selection for yield improvement in Hays Golden. Hallauer (1999) effectively used mass selection methods to adapt five tropical varieties to temperate environments. Carena et al. (2008) were successful in adapting improved varieties to the northern USA (Tables 12.1, 12.2, and 12.3). However, Williams and Davis (1983) were not successful in increasing levels of resistance to southwestern corn borer via mass selection. Effectiveness of selection is determined by the levels of heritability on an individual plant basis. In all cases, improved populations through mass or stratified mass selection methods can continue improvement with recurrent selection methods emphasizing complex traits.

### Table 12.1

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Pollen shedding (days)</th>
<th>Silking (days)</th>
<th>Plant height (cm)</th>
<th>Ear height (cm)</th>
<th>Grain yield (Mg/ha)</th>
<th>Grain moisture (%)</th>
<th>Test weight (Kg/hL)</th>
<th>Cob diameter (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBSK(HI-M)C0</td>
<td>68.1</td>
<td>72.0</td>
<td>201.9</td>
<td>92.0</td>
<td>5.5</td>
<td>28.1</td>
<td>74.4</td>
<td>23.4</td>
</tr>
<tr>
<td>NDBSK(HI-M)C1</td>
<td>67.7</td>
<td>69.6</td>
<td>197.0</td>
<td>91.2</td>
<td>5.4</td>
<td>25.9</td>
<td>74.5</td>
<td>23.4</td>
</tr>
<tr>
<td>NDBSK(HI-M)C2</td>
<td>65.1</td>
<td>69.0</td>
<td>193.0</td>
<td>85.2</td>
<td>5.9</td>
<td>24.0</td>
<td>75.1</td>
<td>22.9</td>
</tr>
<tr>
<td>NDBSK(HI-M)C3</td>
<td>62.5</td>
<td>65.1</td>
<td>190.4</td>
<td>80.7</td>
<td>6.6</td>
<td>21.9</td>
<td>76.3</td>
<td>22.8</td>
</tr>
<tr>
<td>Mean</td>
<td>65.9</td>
<td>68.9</td>
<td>195.6</td>
<td>87.3</td>
<td>5.9</td>
<td>25.0</td>
<td>75.1</td>
<td>23.1</td>
</tr>
<tr>
<td>Average response per year</td>
<td>−1.90</td>
<td>−2.30</td>
<td>−3.83 −3.77</td>
<td>0.37</td>
<td>−2.07</td>
<td>0.63</td>
<td>−0.20</td>
<td></td>
</tr>
</tbody>
</table>

#### 12.4.2 Modified Ear-to-Row Selection

Modified ear-to-row selection is based on among and within half-sib family selection. Because the selection unit is a half-sib family, its theoretical aspects are discussed in Chapters 6 and 7 under half-sib family selection. Here these methods are discussed under the usual nomenclature because of their peculiarities in methodology. Limited data of modified ear-to-row selection (Lonnquist, 1964) indicate this is an effective method of recurrent selection (Paterniani, 1967a, b; Webel and Lonnquist, 1967; Carena and Cross, 2003; Hyrkas and Carena, 2005). Modified ear-to-row has an advantage over mass selection and the original form of ear-to-row selection (Hopkins, 1899) because more than one environment is used to evaluate half-sib families in this particular methodology. Inclusion of more than one environment permits estimation of the genotype by environment interaction term and combining data across environments to determine superior families for recombination. Bias due to environmental forces is removed to a greater extent than for mass selection, but greater resources are needed than for mass selection.
Table 12.2  Means of traits after stratified mass selection for female flowering for NDBS11 maize population cycles across North Dakota environments in 2005, 2006, and 2007 (adapted from Carena et al., 2008)

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Pollen shedding (days)</th>
<th>Silking (days)</th>
<th>Plant height (cm)</th>
<th>Ear height (cm)</th>
<th>Grain yield (Mg/ha)</th>
<th>Grain moisture (%)</th>
<th>Test weight (Kg/hL)</th>
<th>Prolificacy (Ears/plant)</th>
<th>Cob diameter (mm)</th>
<th>Kernel row number</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBS11(FR-M)C0</td>
<td>72.2</td>
<td>73.5</td>
<td>222.3</td>
<td>105.6</td>
<td>5.6</td>
<td>28.0</td>
<td>72.1</td>
<td>1.3</td>
<td>22.5</td>
<td>15.1</td>
</tr>
<tr>
<td>NDBS11(FR-M)C1</td>
<td>70.8</td>
<td>72.5</td>
<td>218.3</td>
<td>103.9</td>
<td>5.8</td>
<td>26.7</td>
<td>72.1</td>
<td>1.2</td>
<td>22.2</td>
<td>14.4</td>
</tr>
<tr>
<td>NDBS11(FR-M)C2</td>
<td>67.9</td>
<td>70.3</td>
<td>206.3</td>
<td>93.6</td>
<td>5.9</td>
<td>24.0</td>
<td>73.5</td>
<td>1.2</td>
<td>22.0</td>
<td>14.4</td>
</tr>
<tr>
<td>NDBS11(FR-M)C3</td>
<td>64.2</td>
<td>67.9</td>
<td>197.9</td>
<td>86.2</td>
<td>6.6</td>
<td>20.8</td>
<td>74.8</td>
<td>1.2</td>
<td>21.7</td>
<td>14.1</td>
</tr>
<tr>
<td>Mean</td>
<td>68.8</td>
<td>71.0</td>
<td>211.2</td>
<td>97.3</td>
<td>6.0</td>
<td>24.9</td>
<td>73.1</td>
<td>1.2</td>
<td>22.1</td>
<td>14.5</td>
</tr>
<tr>
<td>Average response per year</td>
<td>−2.67</td>
<td>−1.90</td>
<td>−8.13</td>
<td>−6.47</td>
<td>0.30</td>
<td>−2.40</td>
<td>0.90</td>
<td>−0.03</td>
<td>−0.27</td>
<td>−0.33</td>
</tr>
</tbody>
</table>
Table 12.3  Means of traits after stratified mass selection for female flowering for NDBS1011 maize population cycles across North Dakota environments in 2005, 2006, and 2007 (adapted from Carena et al., 2008)

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Pollen shedding (days)</th>
<th>Silking (days)</th>
<th>Plant height (cm)</th>
<th>Ear height (cm)</th>
<th>Grain yield (Mg/ha)</th>
<th>Grain moisture (%)</th>
<th>Test weight (Kg/hL)</th>
<th>Ear length (cm)</th>
<th>Ear diam. (mm)</th>
<th>Kernel row number</th>
<th>Prol. (Ears/plt.)</th>
<th>Cob diam. (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBS1011(FR-M)C0</td>
<td>75.9</td>
<td>77.2</td>
<td>213.7</td>
<td>112.9</td>
<td>4.5</td>
<td>32.4</td>
<td>71.8</td>
<td>13.0</td>
<td>38.0</td>
<td>13.7</td>
<td>1.5</td>
<td>21.3</td>
</tr>
<tr>
<td>NDBS1011(FR-M)C1</td>
<td>71.4</td>
<td>73.1</td>
<td>222.3</td>
<td>107.3</td>
<td>5.7</td>
<td>28.7</td>
<td>73.6</td>
<td>14.5</td>
<td>40.4</td>
<td>14.8</td>
<td>1.2</td>
<td>22.6</td>
</tr>
<tr>
<td>NDBS1011(FR-M)C2</td>
<td>69.0</td>
<td>71.6</td>
<td>217.0</td>
<td>105.6</td>
<td>5.9</td>
<td>26.4</td>
<td>74.6</td>
<td>14.6</td>
<td>40.2</td>
<td>15.0</td>
<td>1.2</td>
<td>22.2</td>
</tr>
<tr>
<td>NDBS1011(FR-M)C3</td>
<td>66.0</td>
<td>70.2</td>
<td>211.3</td>
<td>98.7</td>
<td>6.3</td>
<td>24.0</td>
<td>75.6</td>
<td>14.8</td>
<td>40.2</td>
<td>14.3</td>
<td>1.1</td>
<td>22.3</td>
</tr>
<tr>
<td>NDBS1011(FR-M)C4</td>
<td>66.0</td>
<td>69.1</td>
<td>206.6</td>
<td>92.9</td>
<td>6.0</td>
<td>21.5</td>
<td>75.3</td>
<td>14.4</td>
<td>41.1</td>
<td>14.8</td>
<td>1.1</td>
<td>22.2</td>
</tr>
<tr>
<td>MEAN</td>
<td>69.7</td>
<td>72.2</td>
<td>214.2</td>
<td>103.5</td>
<td>5.7</td>
<td>26.6</td>
<td>74.1</td>
<td>14.3</td>
<td>40.0</td>
<td>14.5</td>
<td>1.2</td>
<td>22.1</td>
</tr>
<tr>
<td>Average response per year</td>
<td>−2.48</td>
<td>−2.03</td>
<td>−1.78</td>
<td>−5.00</td>
<td>0.38</td>
<td>−2.73</td>
<td>0.88</td>
<td>0.35</td>
<td>0.78</td>
<td>0.28</td>
<td>−0.10</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Prol (Prolificacy), Diam (Diameter), plt (plant)
Modified ear-to-row selection includes growing a single replication of families in two or more environments. One replication is grown in isolation to facilitate recombination. A sample of ears is taken from the population chosen for improvement. If the half-sib families are to be evaluated in three environments, a single replication of each half-sib family is included in each environment. Separate randomizations are conducted for each environment. The one replication grown in isolation provides data on each half-sib family in addition to recombination. In addition to the half-sib families, the isolation planting includes a bulk of the families that pollinate the half-sib families. The half-sib families are detasseled and cross-pollinated provides seed for each of the progenies. If selection is only on the basis of the half-sib family means for the three environments, the component of variation among family means is $(\frac{1}{8})\hat{\sigma}^2_A$ (Empig et al., 1972). If selections also are made among plants within progeny rows, expected progress includes an additional component of $(\frac{3}{8})\hat{\sigma}^2_A$. Each plant within the progeny rows, for example, can be infected with stalk rot organisms and infested with first brood European corn borers. Data from all environments are used to choose superior families, and individual plant selections are made within selected half-sib families.

Compton and Comstock (1976) offered a further modification using two seasons rather than one for conducting modified ear-to-row selection. All environments in season 1 are used to identify superior families with no isolation planting for recombination. In season 2 only selected families are planted ear-to-row, detasseled, and cross-pollinated with males consisting of a bulk of previously selected half-sib families. Parental control is greater because the selected families are mated with only males formed from a bulk of the selected families. On the other hand, in modified ear-to-row selection conducted in only one season the selected families were mated with males of all families, selected and unselected. Hence, in the prediction equation we have $(\frac{1}{4})\hat{\sigma}^2_A$ for among-family and $(\frac{3}{8})\hat{\sigma}^2_A$ for within-family selection. Compton and Comstock concluded that in temperate zones with only one growing season the two-season scheme would have about 75% as much gain per year as the one-season scheme of modified ear-to-row selection. However, the second season could be grown in winter nurseries.

Modified ear-to-row selection is an intra-population scheme that should be effective in maize populations having adequate genetic variability. It may be used initially to improve a population or after mass selection has been practiced for highly heritable traits. Because of the ever present environmental effects, modified ear-to-row selection would seem more appropriate for improvement of such traits as yield. Because modified ear-to-row is for intra-population improvement, Lonnquist (1964) suggested that two populations exhibiting heterosis in their cross should be submitted to modified ear-to-row selection. The natural procedure would be to use some form of inter-population recurrent selection after response to modified ear-to-row selection seems to be diminishing.

Modified ear-to-row selection does not evolve progenies that are inbred to any extent. Consequently, selected progenies may not be amenable for inbred line development. However, superior progenies can be selfed and incorporated in the common pedigree selection inbreeding system. Modified ear-to-row is also an improvement
over mass selection for identifying superior progenies. With the two-season scheme, a smaller number of selected families can be included in the isolated recombination planting. Selected plants within selected families would be logical candidates for incorporating them in the breeding nursery.

12.4.3 S1 Progeny Recurrent Selection

The S1 progeny recurrent selection has been used for improvement of several traits. Response to S1 progeny recurrent selection usually has been positive (see Chapter 7). It lends itself to evaluation of most traits on a progeny basis. The coefficient for \( \hat{\sigma}^2_A \) is 1 for a restricted genetic model (see Chapters 2 and 6) and an estimate of \( \hat{\sigma}^2_A \) for the population under selection is determined from the component of variance for S1 progenies. Usually, supplies of S1 seed produced on S0 plants are sufficient to permit replicated trials in different environments. Selection among individual S0 plants can be practiced as a basis for self-pollination. Because the heritability on an individual plant basis is usually lower than on a progeny basis in replicated trials, effectiveness of selection will not be as good as estimates of heritability based on replicated progeny trials. Higher heritability estimates of replicated progeny trials were the reason why Penny et al. (1967) and Jinahyon and Russell (1969a) used S1 progeny evaluation rather than S0 plants for selecting individuals having resistance to first brood European corn borer and stalk rot organisms.

Although S1 progeny evaluation lends itself to improvement of most traits of maize, it has not been used as extensively as the half-sib or testcross method for improvement of yield. Types of genetic effects affecting yield expression have caused some hesitation in the use of S1 progeny recurrent selection for yield improvement. Selection based on inbred progenies (S1, S2, etc.) is theoretically more effective for changing frequencies of alleles having additive effects than are testcross methods of selection (see Chapter 6). Testcross methods would be expected to capitalize on those genes that have overdominant effects for the specific tester used.

Selection studies comparing S1 and testcross progeny evaluation are not in complete agreement, but both methods have significantly improved populations under selection. Lonnquist and Lindsey (1964) and Horner et al. (1973) found that S1 and S2 selection were less effective than expected in comparison with testcross selection. Genter and Alexander (1962), Burton et al. (1971), and Genter (1973) found that both methods of selection (S1 and testcross) improved populations for yield of S1 lines, but improvement was greatest with S1 progeny selection. In nearly all instances, inbreeding depression was less for S1 lines progenies extracted from populations improved by S1 progeny selection than for S1 progenies extracted from populations improved by testcross selection. The concerns are how the combining abilities of the populations are affected by the two methods of selection and particularly the possible implications of type of gene action. Present evidence is not conclusive, but it seems that the testcross selection method favors selection of dominant favorable alleles. Burton et al. (1971) found evidence that S1 and testcross
methods of selection either select for different alleles and the difference in their frequencies in selected populations was great enough for the heterosis of the population crosses to increase or genetic drift was great enough to cause differences in allele frequency. Because S₁ and testcross selection emphasize selection for different alleles, differences in allele frequency are to be expected. Tanner and Smith (1987) and Horner et al. (1989) provided further comparisons of inbred progeny vs. half-sib (or testcross) recurrent selection. Tanner and Smith (1987) made the comparisons after eight cycles of selection and reported S₁–S₂ selection to be more effective than half-sib selection after four cycles of selection. However, after eight cycles of selection they found greater response with half-sib selection and no further gains were realized with inbred selection after the fourth cycle. Horner et al. (1989) had similar results in the Fla. 767 population. They interpreted their different responses to selection for overdominant effects which were expressed in the half-sib families founded by crossing with the inbred tester, F6.

Combining ability is a concern, but S₁ progenies selected for recombination to form the next cycle of selection can directly be included in the breeding nursery for further inbreeding and initial testcross evaluations. Jinahyon and Russell (1969b) found that S₁ progeny selection for stalk rot resistance did not change combining ability. It increases average yield level of S₁ progenies, which is desirable for developing lines for use in hybrids. Selected S₁ entries chosen for recombination and included in testcross nurseries to produce testcross seed will be a selected sample of S₁ lines based on S₁ line per se evaluation. All S₁ lines may not be included in testcross trials because some may be discarded in the breeding nursery. Another option would be to include selected S₁ entries only in the breeding nursery and not used in producing testcross seed until a later generation of inbreeding, e.g., S₃ or S₄. The number of selections is usually not great (20–40). Production of hybrid seed in a testcross nursery would not require much additional space. In addition, separate seasons would not be used for the two phases, thus saving time in collecting combining ability information on S₁ lines selected for recombination. Preferences on generation of inbreeding to testcross lines will influence when selected lines are included in the testcross nursery, i.e., S₁ or S₂ vs. S₄ or S₅ (see Chapter 8).

The number of years needed to complete each cycle of S₁ progeny recurrent selection depends on number of growing seasons available and trait under selection (Table 12.4). If selections can be made before pollinations (e.g., first brood European corn borer, seedling vigor, emergence percentage), recombination of selected S₁ progenies can be done in the same season the evaluation trials are conducted. Recombination in the same season will require considerable walking and carrying of pollen to make the necessary balanced crosses among the selected S₁ progenies (e.g., intra-di-allel method, Sezegen and Carena, 2009), but one cycle of selection can be completed each year with the use of off-season nurseries. If it is not desirable to make the crosses the same season, the time required for each generation will be the same as that for traits evaluated later in the growing season (e.g., yield). Two years per cycle is the shortest generation time for traits evaluated after flowering in temperate zones, provided off-season nursery use is possible. Off-season or winter nurseries can be used for recombination and self-pollination to initiate the next cycle of selection. Self-pollinations can be made in season 4 (Table 12.4), but
### Table 12.4 Description of the use of different seasons for conducting S<sub>1</sub> progeny recurrent selection in temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season (Month)</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;In the population under selection (C0), self-pollinate 300–600 S&lt;sub&gt;0&lt;/sub&gt; plants&lt;br&gt;Practice selection among the S&lt;sub&gt;0&lt;/sub&gt; plants at pollination and harvest&lt;br&gt;Harvest S&lt;sub&gt;1&lt;/sub&gt; ears that have sufficient seed for replicated trials</td>
<td>—</td>
</tr>
<tr>
<td>2 (Summer)</td>
<td><strong>Evaluate S&lt;sub&gt;1&lt;/sub&gt; progenies</strong>&lt;br&gt;Conduct replicated trials, select 20–30 S&lt;sub&gt;1&lt;/sub&gt; progenies for recombination&lt;br&gt;If selections can be made before flowering, recombination can also be made this season</td>
<td>—&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>3 (Winter)</td>
<td><strong>Recombine</strong>&lt;br&gt;Recombine the selected S&lt;sub&gt;1&lt;/sub&gt; progenies to form C&lt;sub&gt;1&lt;/sub&gt;&lt;br&gt;If recombination can be made in season 2, plants can be self-pollinated to develop S&lt;sub&gt;1&lt;/sub&gt; progenies for evaluation in season 4</td>
<td>Include selected S&lt;sub&gt;1&lt;/sub&gt; progenies in drought-managed environments and testcross-nurseries for inbreeding and production of testcross seed</td>
</tr>
<tr>
<td>4&lt;sup&gt;b&lt;/sup&gt; (Summer)</td>
<td><strong>Recombine or self-pollinate</strong>&lt;br&gt;Either recombine to form C&lt;sub&gt;1&lt;/sub&gt; Syn 2 or self-pollinate S&lt;sub&gt;0&lt;/sub&gt; plants in the C&lt;sub&gt;1&lt;/sub&gt; Syn 1 to form S&lt;sub&gt;1&lt;/sub&gt; progenies for testing in season 5</td>
<td>Include selected S&lt;sub&gt;1&lt;/sub&gt; progenies in breeding and disease nurseries for selection and inbreeding</td>
</tr>
</tbody>
</table>

<sup>a</sup>S<sub>1</sub> progenies included in replicated trials also can be included in breeding nursery to advance to the S<sub>2</sub> generation

<sup>b</sup>Repeat as described for seasons 1–3

---

Because off-season nurseries usually are not suitable for evaluation trials another generation of recombination can be made before selfing. A decision has to be made as to whether another recombination is more valuable than additional costs of producing S<sub>1</sub> progenies in the off-season. In areas with more than one growing season each year, adjustments can be made for the scheme in Table 12.4 to make maximum use of the available growing seasons. One important way to increase efficiency of selection is to use the least possible number of years to complete each cycle. Length of each cycle depends on the trait under selection and the seasons available to produce progenies for test, evaluation of the progenies, and recombination of selected progenies (see Chapter 7).

Because most breeders are anxious to have new materials in their breeding nurseries, S<sub>1</sub> progeny selection will provide progenies that have been evaluated in one generation and are already selfed. Selections used for recombination in each cycle are logical candidates to include in breeding nurseries for additional selection and testing for inbred line development. Although additional space and time would
be required, all $S_1$ progenies included in evaluation trials could also be included in the breeding nursery for further selection and inbreeding. Some selections not included for recombination also may be included in the breeding nursery. They may be progenies that did not quite meet the standards for recombination but were within the range of the least significant difference of the progenies used. The one generation of $S_1$ progeny testing will screen out poorest progenies, but some that were marginal for including for recombination may be useful after further selection and testing. Dhillon and Khehra (1989) suggested a modification of $S_1$ progeny selection by planting one replication in an isolation to permit intermating which is similar to the modified ear-to-row method. The advantage of the modified method is that it reduces the time frame by one season but disadvantage is that parental control also is reduced.

### 12.4.4 $S_2$ Progeny Recurrent Selection

Mechanics of $S_2$ progeny recurrent selection are similar to those described for $S_1$ progeny selection except that another generation of inbreeding is accomplished before evaluation in replicated trials. The $S_2$ progeny selection has some definite advantages over $S_1$ progeny selection, but an additional year is required to complete each cycle in temperate zones. As in $S_1$ progeny recurrent selection, selection for additive genetic effects is emphasized. The component of variance for among $S_2$ progenies is $(\pi^2)\sigma^2_A$ compared with $\sigma^2_A$ for among $S_1$ progenies for a restricted genetic model (see Chapter 6). Because an extra year is required, strong selection pressure can be applied to larger samples of $S_1$ progenies before conducting $S_2$ progeny trials. The $S_2$ progeny selection usually is used for yield improvement, whereas $S_1$ progeny selection also has been used for traits other than yield.

The sequence of operations for $S_2$ progeny recurrent selection is described in Table 12.5. The population chosen for improvement should have potential as a source for inbred lines with good general combining ability with other elite lines currently used or to be used in the future for hybrid production. The $S_1$ progenies in season 2 should be included in as many nurseries as seed, time, and facilities permit. The primary objective for growing the $S_1$ progeny generation is to eliminate progenies before conducting yield trials in the $S_2$ progeny generation. If sufficient numbers of $S_1$ progenies are available, selection pressure can be increased in discarding susceptible $S_1$ progenies. Use of only one replication in each nursery will not eliminate the possibility of escapes or environmental effects, but if continuous selection pressure is maintained in subsequent cycles of $S_2$ progeny recurrent selection the general level of improvement will increase. Duplicating nurseries across locations and planting dates can improve visual selection. If selections can be made before flowering the number of self-pollinations is reduced, but in most instances selection has to be made after flowering. Selection among and within $S_1$ progenies can be made for several agronomic traits at flowering, e.g., barrenness or poor shoots, tassel type, pollen shed, and any other traits necessary for the line to be useful either as a seed or as a pollen parent in hybrid seed production.
### Table 12.5 Description of the use of different seasons for conducting $S_2$ progeny recurrent selection in temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;In the population under selection (C0), self-pollinate 300–600 $S_0$ plants&lt;br&gt;Practice selection among the $S_0$ plants at pollination and harvest&lt;br&gt;Harvest $S_1$ ears that have sufficient seed for replicated trials</td>
<td>---</td>
</tr>
<tr>
<td>2 (Summer)</td>
<td><strong>Evaluate $S_1$ progenies</strong>&lt;br&gt;Grow the 300–600 $S_1$ progenies in the breeding and pest nurseries. Replicate across locations and planting dates if possible&lt;br&gt;Make selections among- and within-$S_1$ progenies for pest resistance and agronomic traits&lt;br&gt;Advance selected plants to $S_2$ generation by self-pollination&lt;br&gt;Harvest $S_2$ seed on $S_1$ plants that have sufficient seed for replicated yield trials</td>
<td>---</td>
</tr>
<tr>
<td>3 (Summer)</td>
<td><strong>Evaluate $S_2$ progenies</strong>&lt;br&gt;Grow the 100–200 selected $S_2$ progenies in replicated yield trials&lt;br&gt;Select 20–30 progenies for recombination</td>
<td>---$^a$</td>
</tr>
<tr>
<td>4 (Winter)</td>
<td><strong>Recombine</strong>&lt;br&gt;Recombine the selected progenies to form $C_1$&lt;br&gt;Use remnant $S_1$ or $S_2$ seed for recombination</td>
<td>---</td>
</tr>
<tr>
<td>5 (Summer)</td>
<td><strong>Recombine or self-pollinate</strong>&lt;br&gt;Either recombine to form $C_1$ Syn 2 or self-pollinate $S_0$ plants in the $C_1$ Syn 1 to form $S_1$ progenies for testing in season 7</td>
<td>Include selected $S_2$ progenies in the breeding, pest, and/or testcross nurseries for selection, inbreeding, and testcross seed production</td>
</tr>
<tr>
<td>6 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;The off-season nursery is used to produce $S_1$ progenies to initiate next selection cycle</td>
<td>---</td>
</tr>
<tr>
<td>7$^b$ (Summer)</td>
<td><strong>Evaluate $S_1$ progenies</strong>&lt;br&gt;Same as season 2</td>
<td>Include selected progenies in breeding nurseries for further selection and inbreeding&lt;br&gt;Conduct testcross trials in replicated tests</td>
</tr>
</tbody>
</table>

---

$^a$ $S_2$ progenies from yield trials also can be included in the breeding nursery to advance to the $S_3$ generation

$^b$ Repeat as described for seasons 2–5
Major emphasis on selection can be made at harvest for ear size, stalk rot resistance, ear diseases, kernel type, grain filling, fast dry down, etc. One of the needs of S₂ progeny recurrent selection is an ear that has sufficient seed for replicated yield trials. Ear size and number of kernels on the ear also are important for use of lines as parental seed stocks. Replicated yield trials of selected S₂ progenies are conducted in as many environments as seed and facilities permit. Superior yielding progenies are selected for recombination based on the combined data. The S₂ progenies included in yield trials can also be included in the breeding nursery for additional selection and selfing. By the time yield trials are completed, S₃ seed of selected progenies will be available for inbred line development which is a major advantage. Seed supplies may be a limiting factor if remnant S₂ seed is used for recombination.

Mechanics of recombination are the same as for S₁ progeny recurrent selection. The breeder has a choice, however, of which generation of seed to use for recombination, whether to use remnant S₁ or S₂ generation seed. Effects of inbreeding are greater with S₂ seed than with S₁ seed (see Chapter 9) but additional selection within S₁ progeny rows was imposed in selecting S₂ progenies for yield traits. Therefore, using remnant S₂ seed takes advantage of it. If sufficient S₂ remnant seed is not available or one is concerned about effects of inbreeding, remnant S₁ seed can be used for recombination. As an example, if 20 S₁ progenies are recombined vs. 20 S₂ progenies, effects of inbreeding are 22% for S₁ progenies vs. 31% for S₂ progenies in the fifth cycle of recurrent selection. Or it would require 18 S₂ progenies vs. 12 S₁ progenies to have the same level of inbreeding. If remnant S₂ generation seed is used for recombination, it is an advantage to have another generation of recombination in season 4 (Table 12.4). Because season 6 will generally not be as satisfactory for evaluating S₁ progenies for temperate zones, the additional generation of recombination will not increase the time for each cycle of recurrent selection. The S₁ progenies can be obtained in the next off-season nursery (season 6, Table 12.4).

Practical usefulness of S₂ progeny recurrent selection is to include selected S₂ progenies in the breeding nurseries for selection, inbreeding, and producing testcross seed. The S₂ progenies selected for recombination are the survivors of intensive multiple trait selection in the S₁ and S₂ generations and should have potential in developing new lines. They are an elite sample of progenies that have to be yield tested in hybrid combinations to determine their combining ability with other elite lines in the S₂ and later generations. The S₂ progeny selection method requires at least 3 years for each cycle. For an applied breeding program in search of new lines, S₂ progeny recurrent selection seems to be a good choice. The increased length of cycle has advantages because of increased opportunities to practice selection before yield testing and because lines are already at an advanced stage of inbreeding. Selection in the S₁ generation eliminates progenies that do not meet accepted standards. The important point is to recombine selected progenies to continue selection for future cycles. A logical division of germplasm in the heterotic response of elite lines, such as dent vs. flint and BSSS vs. Lancaster, suggests S₂ progeny recurrent selection in two populations to test combining ability of selected progenies, i.e., testing selections from a dent population against those from a flint population.
12.4.5 **Half-Sib Recurrent Selection**

Half-sib recurrent selection includes several versions depending on types of testers used to produce the half-sib (testcross) progenies for evaluation (see Chapters 7 and 8). Half-sib recurrent selection has been useful for cultivar development programs as B73 was developed through the integration of half-sib recurrent selection and the pedigree selection method of inbred line development. Mechanics of conducting half-sib recurrent selection are similar for different types of testers. The choice of testers depends on the breeder’s current knowledge of relative importance of different types of gene action, materials available, and stage of breeding program. In contrast to S₁ and S₂ progeny recurrent selection based on performance of progenies themselves, half-sib recurrent selection of progenies is based primarily on their performance in crosses with other genotypes (combining ability). Recurrent selection on the basis of S₁ and S₂ progenies emphasizes selection that changes frequencies of genes having additive effects more than do recurrent selection methods involving testcrosses (see Chapter 6). If overdominance is important at some loci, half-sib recurrent selection would favor selection for alleles that have overdominant effects more than does S₁ and S₂ progeny evaluation. Stage of testing also will influence the decision for use of half-sib recurrent selection, which emphasizes early testing for combining ability. Comparisons of inbred progeny evaluation and half-sib progeny evaluation are discussed in Chapter 8.

Half-sib progeny evaluation can be conducted by producing either S₀ plants or S₁ plant testcrosses, but it seems that use of S₁ plants would be preferable for inbred line development. Description of half-sib selection in Table 12.6 is based on testcrosses from use of S₁ plants. This procedure requires one additional season, but the number of years required to complete one cycle of recurrent selection does not change. Its major advantage is that S₁ progenies can be screened in choosing the S₁ plants used to produce the testcrosses. The S₁ progenies represent the S₀ plant genotype and may be included in different nurseries and selections made at either pollination or harvest for plants to include in testcross trials. Testcrosses of only selected plants in selected S₁ progenies are yield tested. Off-season nurseries can be used to produce S₁ progenies without increasing the cycle interval for half-sib recurrent selection.

The type of tester will influence the amount of effort required to produce testcross seed. Use of either an inbred or a single-cross tester will require making only enough crosses to have sufficient seed supplies for replicated yield trials. For a single-cross tester, perhaps only two or three ears are necessary. Use of the parental population or some other type of population with a broad genetic base involves some sampling problems. To have the gametic array of the tester adequately represented in the testcross, 6–10 plants of the tester should be pollinated (Sprague, 1939; Noble, 1966). The number of plants pollinated with use of a broad genetic base tester, however, may not be any greater than with use of an inbred tester. We need adequate sampling when a broad genetic base tester is used, and we need to ensure an adequate quantity of seed when an inbred tester is used. The number of pollinations required, therefore, may not be greatly different for different testers.
Table 12.6 Description of the use of different seasons for conducting half-sib recurrent selection based on S1 testcrosses for temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;In the population under selection (C0), self-pollinate 300–600 S0 plants&lt;br&gt;Practice selection among the S0 plants at pollination and harvest</td>
<td></td>
</tr>
<tr>
<td>2 (Summer)</td>
<td><strong>Produce testcrosses</strong>&lt;br&gt;Grow 300–600 S1 progenies in the breeding and pest nurseries&lt;br&gt;Make selections among S1 progenies when possible before pollination&lt;br&gt;Produce testcrosses (or half-sibs) of selected S1 plants in selected S1 progenies&lt;br&gt;Self-pollinated the S1 plants used to produce the testcrosses</td>
<td></td>
</tr>
<tr>
<td>3 (Summer)</td>
<td><strong>Evaluate testcrosses</strong>&lt;br&gt;Grow 100–200 testcrosses in replicated yield trials&lt;br&gt;Select 20–30 progenies for recombination</td>
<td>—a</td>
</tr>
<tr>
<td>4 (Winter)</td>
<td><strong>Recombine</strong>&lt;br&gt;Recombine either selfed (S1 or S2) progeny or remnant half-sib seed to form C1</td>
<td></td>
</tr>
<tr>
<td>5 (Summer)</td>
<td><strong>Recombine or self-pollinate</strong>&lt;br&gt;Either recombine to form C1 Syn 2 or self-pollinate S0 plants in the C1 Syn 1 to develop S1 progenies for selection and testcrossing in season 7</td>
<td>Include S2 progenies or selected half-sibs in the breeding, pest, and/or testcross nurseries for selection, inbreeding, and testcross seed production</td>
</tr>
<tr>
<td>6 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;The off-season nursery is used to produce S1 progenies to initiate next selection cycle</td>
<td></td>
</tr>
<tr>
<td>7b (Summer)</td>
<td><strong>Produce testcrosses</strong>&lt;br&gt;Same as season 2</td>
<td>Include selected progenies in breeding nurseries for further selection and inbreeding&lt;br&gt;Conduct testcross trials in replicated tests</td>
</tr>
</tbody>
</table>

---

a S2 progenies of the S1 plants used to produce testcrosses also can be included in the breeding nursery to advance to the S3 generation

b Repeat as described for seasons 2–5
Careful planning of the nursery is necessary in half-sib recurrent selection. In pest nurseries the replicated or non-replicated S1 progenies may be grown in sequence. In breeding nurseries for half-sib recurrent selection, it is necessary to alternate S1 progeny rows or S1 progeny ranges with the tester parent. To minimize the mechanics and time for producing testcrosses, the tester parent should be close to the S1 progenies. Planting the tester parent in ranges adjacent to each range of S1 progenies minimizes time and labeling required to make the testcrosses. The tester parent and S1 progenies may have poor timing at flowering because of inbreeding depression and maturity of S1 progenies. It may be necessary to have one or two delay plantings of the tester parent. The first planting is made at the same time as the S1 progenies with subsequent plantings made 5–10 days later. Lower planting densities are used for the first planting to permit row identification for delayed plantings. If an inbred tester is used it may be necessary to delay plant the S1 progenies, but this usually is not critical. If the tester and S1 progenies are planted end-to-end in different ranges, harvesting also is simplified. For instance, the NDSU maize breeding program plants, for this purpose, a range of males (S1 lines), a range of females (industry tester), and a second range of males (S1 lines). The middle range (females) is divided in to two and the two ranges of males are used for selfing and crossing. Pair crosses between S1 progenies and industry testers have also worked out for producing half-sib testcrosses. Testcrosses produced on inbred and single-cross testers are shelled in bulk, but equal quantities of seed to meet the requirements for testing are shelled from each ear when a broad genetic base tester is used (e.g., if 500 seeds required for yield testing and eight testcrossed ears are available, 64 seeds are shelled from each ear). The half-sib progenies are evaluated in replicated trials in three or four environments in season 3 (Table 12.6).

Selected progenies to be recombined to form the next cycle population are determined by performance of half-sib progenies. Remnant S1 or S2 seed of plants used to produce the testcrosses are used for recombination. Recombination can be completed in off-season nurseries (winter for temperate zones), with either recombination or self-pollination the following season. Since another off-season nursery is available for temperate zones an additional generation of recombination (season 5, Table 12.6) seems desirable, particularly if S2 seed is used for recombination.

Half-sib progeny recurrent selection also is amenable to generating elite and diverse material for pedigree selection. The S1 plants used in making the testcrosses also are self-pollinated to produce S2 seed. Remnant S2 seed can be included in the inbred line development process after the testcross information has become available with the advantage of having early-generation hybrid testing already initiated. Selected S2 progenies have had selection in the S0 and S1 generations and testcross information in the S1 generation; thus they have had considerable selection pressure as inbred progenies and for combining ability with the tester used to produce the half-sib progenies. Or if S2 generation seed was included in the breeding nursery in season 3, S3 generation progenies are available. The additional information on combining ability was not available for the S1 and S2 progeny recurrent selection methods discussed previously. If the breeder considers early testing of S1 plants an adequate test for combining ability, half-sib recurrent selection would be appealing.
Again, some additional S₂ or S₃ progenies may be included in the breeding and testcross nurseries that were not included in 20–30 selections recombined for the next cycle population. Additional inbreeding, selection, and testcrossing may reveal some superior lines that were not ranked correctly on the basis of initial testcrosses, especially due to genotype by environment (e.g., year) interactions.

The sequence of methods for half-sib recurrent selection in Table 12.6 was described for only one population. An applied breeding program for development of new inbred lines and hybrids will include some populations that exhibit a greater heterotic response than others. A natural sequence is to pair populations that have the greatest heterotic response and potential usefulness as fruitful breeding populations and use one as the tester for the other, and vice versa. The description of half-sib recurrent selection in Table 12.6 would be applicable, but rather than one population there would be two populations under recurrent selection. To reduce the amount of testing in each season, it may be advisable to alternate the sequence of breeding activities for the two populations. Testing in alternate seasons would reduce the number of test plots, particularly when considered in reference to other testing priorities of the breeding program. If each population is the tester for the other, recurrent selection procedures would be similar to reciprocal recurrent selection described by Comstock et al. (1949). Use of an inbred line or a single-cross hybrid as tester would be equivalent to the method of recurrent selection described by Russell and Eberhart (1975). Comstock (1979), however, showed theoretically, that the advantages of using either an inbred or a single cross as testers, as suggested by Russell and Eberhart (1975), depends on the allele frequencies of the tester. For long-term programs, use of population as tester would be preferable.

Half-sib recurrent selection conducted concurrently in two breeding populations provides a natural partition of the derived lines for testing in hybrids. It is logical that lines extracted from a Reid/BSSS or dent-type population would be tested with lines extracted from Lancaster Surecrop/non-BSSS or flint-type populations. Selections obtained from the two half-sib recurrent selection programs would generate lines for the breeding nursery for both sides of the heterotic pedigree, which is highly desirable. If one highly promising population is used for half-sib recurrent selection, the lines extracted may not have a natural classification as to type of tester needed for further evaluation or how they should be used in currently produced hybrids. This dilemma is not serious if the population under half-sib recurrent selection has had previous testing and the heterotic response is known with reference to other breeding materials. Moreover, if lines are developed from genetically diverse populations they may have the advantage to combine well across heterotic groups (Barata and Carena, 2006). However, the disadvantage would be to increase the number of testers and hybrid combinations.

12.4.6 Full-Sib Family Recurrent Selection

Full-sib family recurrent selection in the Jarvis and Indian Chief varieties was reported by Moll and Stuber (1971). Comparisons were made between full-sib family recurrent selection (intra-population) and reciprocal recurrent selection (RRS)
Responses of both varieties to full-sib family recurrent selection were 2.1 times greater than their response to RRS. Hence full-sib family recurrent selection was more effective than RRS for improvement of the populations themselves. In contrast RRS was 1.3 times more effective than full-sib family recurrent selection for improvement of the variety cross between Jarvis and Indian Chief. Results from this study seem to agree with the basic principles of the two recurrent selection schemes because RRS was designed primarily for improvement of the cross of two populations. Full-sib family recurrent selection was described by Mather (1949) as bi-parental crosses.

Unlike producing inbred or half-sib progenies the mechanics for producing full-sib families require more details that are identified in tassel bags. The number of nursery rows needed among methods for a population is similar, not large (e.g., 20–40 rows of 30–40 plants per row per population). In this case, it is advisable to tag not only end plants but also plants within rows to know the exact location of the plants being crossed. Full-sib families are established in populations chosen for improvement by crossing one plant with another. Reciprocal crosses can be made between the two plants included in a cross and the seed bulked (often from two ears) for evaluation trials. Each of these pairs would be a full-sib family. The number of full-sib families deemed sufficient for adequate sampling (e.g., 200 progenies produced from 400 plants) of the population under improvement would be produced. Selection of $S_0$ families to include for producing full-sib families can be made at pollination and harvest (e.g., discard full-sib pairs from plants with weak stalks). Selection at pollination would be only for obvious phenotypic plant traits (e.g., pollen shed, shoot emergence, plant type), but selection at harvest could be made for ear type, seed set, and stalk quality. Adequate crosses would be needed to allow for discarding at harvest. Selection would not be as effective as on progenies that have been inbred and can be replicated (e.g., $S_1$ or $S_2$). Full-sib families are evaluated in replicated trials and superior full-sib families identified. Remnant full-sib seed of superior families are recombined to initiate the next cycle of selection (Table 12.7).

At least three seasons are required to complete one cycle of full-sib family recurrent selection. In temperate zones it requires 2 years to complete one cycle to

1. produce full-sib families,
2. evaluate full-sib families, and
3. recombine selected full-sib families, which is the off-season or winter nursery.

If full-sib families are produced in the winter nursery, five seasons in 2 years are used to complete each cycle with two generations (or seasons) of recombination included. If intercrosses produced during recombination are full-sibs, they could be included in evaluation trials to complete 1 cycle per year.

Full-sib family recurrent selection does not lend itself to generating progenies quickly for inbred line development. Remnant full-sib seed is available, but the families are non-inbred at this stage. The selected full-sib families are produced from non-inbred progenies with only mass selection of $S_0$ plants. Lines could be derived from selected full-sib families but it may be preferable to extract lines from the population improved by full-sib recurrent selection, which is a procedure similar to
Table 12.7 Description of the use of different seasons for conducting full-sib recurrent selection in temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td>Produce full-sib family progenies</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>In the population under selection (C0), produce 150–200 full-sib crosses by reciprocal pollinations of two selected plants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Practice selection among S0 plants at pollination and harvest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Harvest full-sib crosses that have sufficient seed for replicated yield trials</td>
<td></td>
</tr>
<tr>
<td>2 (Summer)</td>
<td>Evaluate full-sib progenies</td>
<td>—^a</td>
</tr>
<tr>
<td></td>
<td>Conduct replicated trials of 100–200 full-sib progenies at 3–4 locations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Select (e.g., index selection for multiple traits) 20–30 top full-sib progenies for recombination</td>
<td></td>
</tr>
<tr>
<td>3 (Winter)</td>
<td>Recombine</td>
<td>Include selected full-sib progenies in winter nurseries for inbreeding, drought screening, and testcrossing</td>
</tr>
<tr>
<td></td>
<td>Recombine the selected progenies to form C1. The bulk-entry method of recombination offers the least number of rows (e.g., two times the number of progenies selected)</td>
<td></td>
</tr>
<tr>
<td>4 (Summer)</td>
<td>Recombine or produce full-sib progenies</td>
<td>Include selected S1 progenies in breeding and disease nurseries for selection and inbreeding</td>
</tr>
<tr>
<td></td>
<td>An additional recombination can be obtained to form C1 Syn 2 population or full-sib progenies of the C1 Syn 1 population can be produced for testing in season 6</td>
<td></td>
</tr>
<tr>
<td>5b (Winter)</td>
<td>Produce full-sib progenies</td>
<td>Include selected S2 progenies in winter nurseries for inbreeding, drought, and testcrossing</td>
</tr>
<tr>
<td></td>
<td>The next cycle of selection is initiated in the C1 Syn 2 population by producing the full-sib crosses</td>
<td></td>
</tr>
</tbody>
</table>

^a Full-sib progenies can also be included in the breeding nursery for selfing (Selfing can be conducted while recombining with the bulk-entry method). Selfed progenies can be used for recombination and advanced in the breeding nursery (Sprague and Eberhart, 1977)

^b Repeat as described for seasons 1–4

that described for mass selection and modified ear-to-row selection. Full-sib family recurrent selection can be modified to make it more amenable to generating lines for pedigree selection. Either S0 plants can be self-pollinated to produce S1 progenies or S0 plants that produce seed on more than one ear can be used to produce full-sib families for evaluation. Sprague and Eberhart (1977) suggested that full-sib families be included in the breeding nursery the same season of yield trials and selfed to generate S1 progenies for further evaluation. However, it may be preferred to wait
for full-sib progeny data to only include selected ones. Modifications are discussed in more detail in reciprocal full-sib selection.

Full-sib family recurrent selection is an intra-population improvement scheme, but two populations that have manifested heterosis in the population cross can be chosen for full-sib family recurrent selection. Although it seems that intra-population full-sib recurrent selection is not as effective as RRS for improvement of the population cross, comparison of improvements shows that full-sib family recurrent selection may be a logical choice for improvement of two populations as source populations for extraction of lines. Moll and Stuber (1971) showed that full-sib selection seems to be a viable recurrent selection scheme for improvement of breeding populations. Minor modifications to facilitate extraction of lines would increase potential of full-sib family recurrent selection. Several populations and lines developed at CIMMYT and NDSU have extensively utilized this methodology.

12.5 Inter-Population Genetic Improvement

12.5.1 Reciprocal Recurrent Selection

RRS was recommended by Comstock et al. (1949) for improvement of performance of the cross between two populations that included selection for specific and general combining abilities. Hull (1945) believed that selection was not effective in maize because of the paucity of additive genetic effects and that overdominant loci played an important role in manifestation of heterosis in crosses of inbred lines. Hull suggested that to maximize effectiveness of selection for overdominant loci either an inbred line or a single-cross hybrid should be used as tester. For improvement of a breeding population the narrow genetic base tester should be used in each cycle of selection, hence, recurrent selection for specific combining ability. It was shown in Chapter 5; however, that adequate additive genetic variability was present in maize populations to expect progress from selection based on additive genetic effects with partial to complete dominance. Some earlier estimates obtained in F_2 populations suggested overdominance, but subsequent studies in random mated F_2 populations showed that estimates of overdominance were largely due to repulsion phase linkages (see Chapter 5). Pseudo-overdominance was detected because of linkage biases in the F_2 populations produced from inbred lines. Recurrent selection for specific combining ability was in contrast to selection for general combining ability, which was based primarily on additive genetic effects (Jenkins 1940). Because of divergence of opinion on whether selection should be emphasized for general or for specific combining ability, Comstock et al. (1949) proposed the use of RRS, which would include selection for both general and specific combining abilities. If the populations under RRS were to have value in breeding programs, either inbred lines extracted from the two populations or the population cross can be used to produce hybrids. Evidence summarized in Chapter 7 and by Hallauer and Carena (2009) shows that RRS effectively improves the population crosses (direct responses) with smaller changes (indirect responses) in the population themselves.
Comstock et al. (1949) showed that if overdominant loci are more important, recurrent selection for specific combining ability will be more effective than RRS. If only additive genetic effects with partial to complete dominance are more important, recurrent selection for general combining ability will be more effective than RRS. But if both types of gene effects are operative in expression of a trait, RRS will be the more effective method of selection. The advantages of recurrent selection for specific and general combining abilities over RRS for the specific cases were not great. Hence it seemed that RRS would be at least as effective as recurrent selection for the specific cases and more effective if both types of gene action were involved. If additive genetic effects with partial to complete dominance are of primary importance, most of the recurrent selection schemes should be equally effective. This is evidenced by the summaries in Chapter 7 and by Sprague and Eberhart (1977), Hallauer et al. (1988), Hallauer (1992), and Hallauer and Carena (2009). Recurrent selection for specific combining ability, as discussed previously, also seems to be effective for general combining ability (e.g., Horner et al., 1973; Walejko and Russell, 1977; Hoegemeyer and Hallauer, 1976). Hence the usefulness of RRS may seem limited. But it is a useful selection scheme for breeding programs, particularly if one is interested in developing new lines from two breeding populations that manifest heterosis in their crosses. Inter-population selection has also an advantage over intra-population selection if there are multiple alleles unless one can accumulate all the alleles in one population. Accumulation of all alleles in one population, however, negates heterosis between populations. It may take several cycles of selection in the hybrid population to obtain intra-population hybrids that are equivalent to hybrids obtained from use of inter-population selection (i.e., RRS), although Hallauer and Carena (2009) have shown exceptions. Moreover, Carena and Wicks III (2006) have shown heterotic combinations could be exploited without the need for long-term RRS programs. Overall, inclusion of populations should be determined on the basis of the heterotic response of their crosses. In most cases, a natural method is provided for selecting and testing lines derived by RRS. Information from North Carolina, Iowa, and North Dakota RRS programs demonstrate that selected lines from improved populations perform better in hybrids than those from original populations (Russell and Eberhart, 1975; Moll et al., 1977; Betran and Hallauer, 1996; Carena et al., 2009a).

RRS requires 3 years to complete each cycle in temperate zones that have winter nurseries for making self-pollinations and recombination (Table 12.8). Description of RRS in Table 12.8 used S₁ plants to make the testcrosses. Either S₀ or S₁ plants can be used, but additional selection can be made among S₁ progenies for other traits that are important. The S₁ progenies can be developed in the winter breeding nursery and do not extend the length of the cycle. Selection can be made among S₀ plants at pollination and harvest for advancing to the S₁ generation. The S₁ progenies are planted in the breeding and pest nurseries and selection is applied for traits as discussed for intra-population schemes. Selected plants within the selected S₁ progenies are crossed with the opposite parental population, which is the tester population. That is, S₁ plants from population A (e.g., Iowa Stiff Stalk Synthetic) are crossed with population B (e.g., Iowa Corn Borer Synthetic No. 1), the tester population for population A. Also S₁ plants from population B are crossed with
Table 12.8 Description of the use of different seasons for conducting half-sib reciprocal recurrent selection for temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;In the two populations (C0), self-pollinate 300–600 S0 plants in each population&lt;br&gt;Practice selection among the S0 plants at pollination and harvest</td>
<td>—</td>
</tr>
<tr>
<td>2 (Summer)</td>
<td><strong>Produce testcrosses</strong>&lt;br&gt;Grow the 300–600 S1 progenies in the breeding and pest nurseries&lt;br&gt;Use selected S1 progenies and make testcrosses with the opposite population. Cross each selected S1 plant with 6–10 plants. Also, self-pollinate each S1 plant used in testcrosses (e.g., 100–200)</td>
<td>—</td>
</tr>
<tr>
<td>3 (Summer)</td>
<td><strong>Evaluate testcrosses</strong>&lt;br&gt;Grow 100–200 testcrosses in replicated yield trials and select top 20–30 selfed progenies for recombination. There will be two sets of testcrosses, one for each of the two populations under selection</td>
<td>— a</td>
</tr>
<tr>
<td>4 (Winter)</td>
<td><strong>Recombine</strong>&lt;br&gt;Recombine the selfed progenies of the selected testcrosses to form C1 Syn 1</td>
<td>Include selfed progenies in drought-managed nurseries for screening and inbreeding</td>
</tr>
<tr>
<td>5 (Summer)</td>
<td><strong>Recombine</strong>&lt;br&gt;Recombine C1 Syn 1 to form C1 Syn 2</td>
<td>Grow S2 or S3 progenies of selfed testcrosses in breeding and testcross nurseries. Practice additional selection within and among S2 progenies</td>
</tr>
<tr>
<td>6 (Winter)</td>
<td><strong>Self-pollinate</strong>&lt;br&gt;In the two C1 populations, self-pollinate 300–600 S0 plants in each population&lt;br&gt;Practice selection among the S0 plants at pollination and harvest</td>
<td>Include selfed progenies in drought-managed nurseries for screening and inbreeding</td>
</tr>
<tr>
<td>7 (Summer)</td>
<td>Repeat as in seasons 1–5</td>
<td>Grow selected S3 or S4 progenies in breeding nursery and testcrosses of selected S2 progenies in replicated tests</td>
</tr>
</tbody>
</table>

a The S2 progenies can be grown in the breeding nursery the same season the testcrosses are grown. Selection among and within S2 progenies can be made and advanced to the S3 generations
population A, the tester for population B. The number of plants that the S₁ plants are crossed with should be 6–10. Each S₁ plant used to produce a testcross also is self-pollinated to produce S₂ generation seed. All ears are harvested for each S₁ testcross. Equal quantities from each harvested ear are needed to form an adequate seed lot for yield trials. If two isolation fields (at least 200 m from other maize) are available, the S₁ progenies (population A) can be planted ear-to-row, detasseled, and pollinated by males of population B. Similarly, S₁ progenies of populations B can be included in the second isolation to be pollinated by population A. Because two populations are undergoing selection, two sets of testcrosses are included in the yield trials. Therefore, the number of entries for yield trials will be two times that for intra-population improvement, but if two populations are undergoing intra-population recurrent selection the number of yield test plots will not be different. The two sets of testcrosses are grown in separate experiments. The S₂ generation progenies of the S₁ testcrosses can also be included in the breeding nursery to advance to the S₃ generation. After harvesting, data are summarized from all locations and highest yielding testcrosses are identified for each of the two sets of testcrosses. Because intra-population half-sib progenies are evaluated, S₁ or S₂ seed is used for recombination. Parental control is greater with S₁ and S₂ seed than with remnant half-sib seed (see Chapter 6). Top progenies from each population need to be recombined. Therefore, two recombination sets are needed and can be made in the winter nursery following harvest. Pollination efforts can be reduced by detasseling all plants being used as either male or female while discarding shoot-bags from plants used while making the sampling or gametes more representative. The S₁ progenies can be developed the following summer, but another generation of recombination is desirable, particularly if S₂ seed is used for recombination. The S₁ progenies can be developed the following winter season (season 6, Table 12.8) and not increase the cycle time. After obtaining S₁ progenies from the C₁ cycle, the process is repeated in subsequent cycles.

Two sets of selected S₂ progenies (or S₃ progenies if S₂ progenies were grown in season 3) are available to include in the breeding nursery in season 5. The S₂ progenies are the survivors of plants from S₁ progenies that were selected to make the testcrosses and had the highest yielding testcrosses. Hence the S₂ progenies were selected from S₁ progenies that had desirable agronomic traits and above average general combining ability. Additional selection among S₂ progenies can be made either before or at the same time the lines are included in the testcross nursery. The stage of additional testing depends on the preference of the breeder, but testcross seed can be produced the same season without much additional expense. Whether testcross seed is included in testcross trials depends on whether S₂ progenies were deemed desirable for continuing in the breeding nursery.

The logical tester for the two sets of testcrosses again depends on the previous heterotic pattern of the germplasm. For example, the tester for the lines from Iowa Stiff Stalk Synthetic will be one of Lancaster Surecrop origin (e.g., Mo17), and the tester for the lines from Iowa Corn Borer Synthetic No. 1 will be one of Iowa Stiff Stalk Synthetic (Reid) origin (e.g., B73). However, it is common, for instance, in the NDSU maize breeding program, to use current industry testers with
several events (TR1017, TR3030, TR1914, TR3622) to evaluate the genetic potential of lines in both early- and late-generation testing. Rather than using a common tester for each set of lines, testcrosses can also be produced by use of the cross-classification (design II) mating scheme. If, for example, 20 lines are selected from the A and B populations, 5 lines from A can be crossed with 5 lines from B to produce 25 testcrosses. This procedure would require four sets of design II crosses or 100 testcrosses. Use of design II requires additional effort at pollination and more testcrosses are included in testcross trials. Use of a common tester for the two sets of selected S2 lines would result in only 40 entries, whereas the example for the design II scheme would have 100 entries. The design II scheme is used only when no logical tester is available for selected lines. Otherwise, it is easier and cheaper to include selected lines in the testcross nurseries. Partial diallel crosses between selections also could be made and BLUP analyses used to predict crosses that were not included for evaluation (Bernardo, 1996).

Although RRS was the first inter-population scheme proposed for the cyclic improvement of two breeding populations, the scheme has not been tested as much as desirable. Only two RRS programs were being conducted in the USA:

1. In North Carolina for the two open-pollinated varieties, Jarvis and Indian Chief, which have been discontinued.
2. In Iowa for the two synthetic varieties, Iowa Stiff Stalk Synthetic and Iowa Corn Borer Synthetic No. 1, which continued to 17 cycles of RRS completed.

Both programs have been conducted for 50–60 years and significant improvements have been made in the population cross for both programs with limited evidence that genetic variability had decreased (Moll et al., 1977; see Chapter 5).

Two programs also were being conducted in Africa:

1. At Kitale, Kenya, in KII and Ec573 (Darrah et al., 1978).
2. At Pietermaritzburg, South Africa, in Teko Yellow and Natal Yellow Horsetooth (Gevers, 1975).

Paterniani and Vencovsky (1978) reported 3.5% gain from modified RRS in Dent Composite and Flint Composite in Brazil. Three cycles of selection have been completed for the African and the Brazilian programs with use of modified RRS.

RRS seems to be a logical recurrent selection scheme for the development of new lines, provided the selection program is combined with cultivar development. It provides for the development of new lines for both sides of the hybrid pedigree, e.g., Reid vs. Lancaster Surecrop and flint vs. dent. Its limited use may be due to its seemingly more complex than intra-population schemes, but if one conducts intra-population improvement in two populations as suggested previously, RRS is no more complex than half-sib selection in two populations. Hallauer and Carena (2009) summarized the responses reported for the RRS programs and evidence that RRS was effectively changing the allele frequencies because heterosis increased with continued selection. It seems that RRS should be prominent in applied breeding programs to develop new, unique genotypes that could provide genetic variation.
in pedigree selection programs. But because of their long-term nature, their seemingly complex methods, and the resources required to conduct them, RRS methods are presently receiving very limited attention.

### 12.5.2 RRS Based on Testcrosses of Half-Sib Families

RRS based on testcrosses of half-sib families, a modification of the original RRS, was suggested by Paterniani (1967b) and reduces effort for making testcrosses because they are obtained by open pollination in isolated fields. The main genetic difference from the original procedure is the type of parentage among individuals in testcrosses (selection units). Individuals are related as half-sibs and cousins, whereas in the original RRS individuals are related as half-sibs among female–male progenies and as full-sibs within female progenies.

A cycle of the modified method starts with 200 or more open-pollinated ears (half-sib families) from population A that are planted ear-to-row as females in an isolation block. Male rows are planted with seeds of population B with ratio of 3:1, 2:1, or 4:2 (female:male). In a separate isolation block a similar number (200 or more) of open-pollinated ears (half-sib families) of population B is planted as females and population A is used as male. It may be desirable to plant a greater number of half-sib families and discard undesirable ones at harvest. The second phase is evaluation of testcrosses (half-sib family A × population B and half-sib family B × population A) in replicated yield trials. Selection is on the basis of yield means and agronomic traits. Recombination is performed in populations A and B for the selected half-sib family progenies using remnant seeds from the selected half-sib families (see Table 12.9).

Paterniani and Vencovsky (1977) evaluated this type of selection methodology and found a significant response to selection in the population cross (7.5%), although (as emphasized by the authors) genetic properties of the populations may have been an important factor in the response to selection. Simplicity of the scheme and enough seed supplies for testing were cited as advantages. In addition, the scheme may be more appropriate for long-term objectives because of the larger number of progenies tested and the nature of the recombination unit. The method uses a smaller portion of the genetic variability but it permits a higher selection pressure because effective population size is about four times the number of families tested. Therefore, the breeder can easily protect the populations against low effective sizes or depletion of genetic variability due to genetic drift during selection (Paterniani and Vencovsky, 1977).

### 12.5.3 RRS Based on Half-Sib Progenies of Prolific Plants

RRS based on half-sib progenies of prolific plants is another modification of the original RRS. The main difference between this modification and RRS is that individuals within the selection units (testcrosses) are related only as half-sibs and the
Table 12.9 Description of the method of reciprocal recurrent selection based on testcross of half-sib families and its integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Winter)</td>
<td>\textit{Testcross production in isolation}\nEars from isolated open-pollinated fields (A and B) are planted ear-to-row as females in detasseling blocks where the opposite population is used as male to produce the testcrosses</td>
<td>—</td>
</tr>
<tr>
<td>2 (Summer)</td>
<td>\textit{Evaluate testcrosses}\nConduct replicated trials of 100–200 testcrosses (A × B and B × A)\nIdentify the best testcrosses for both half-sib tests</td>
<td>—</td>
</tr>
<tr>
<td>3 (Winter)</td>
<td>\textit{Recombine}\nRemnant seed of half-sib families corresponding to the best testcrosses are planted for recombination for both sets. The diallel mating design can be used but a simpler procedure is the use of isolation blocks where the selected progenies are used as females (detasseled) and a bulk of them as males</td>
<td>Include selected half-sib progenies in drought-managed environments and testcross nurseries for inbreeding and production of testcross seed</td>
</tr>
<tr>
<td>4 (Summer)</td>
<td>Repeat as in season 1–3</td>
<td>Include selected S₁ progenies in breeding and disease nurseries for selection and inbreeding</td>
</tr>
</tbody>
</table>

recombination unit is a half-sib family instead of an S₁ family. In the first phase testcross progenies are obtained in isolated detasseling blocks. In one isolation block, population A is planted as females (detasseled rows) and population B as males and vice versa in the other. Female rows within isolations are planted as either bulked seed or ear-to-row. The second (lower) ear shoots of prolific plants in the female rows in each block are covered. The first (upper) ear of each plant is open pollinated by the opposite population (male) in each block. When the tassels start pollen shedding, a sample of pollen is collected each day from the male rows in each field. The bulk pollen samples are used to pollinate the second ears of plants from the same population (female rows) in the other field. Therefore, pollen from male A plants in the first field is used to pollinate female A plants in the second field and vice versa. Pollen from male rows is collected each day from at least 50 plants and passed through a sieve to eliminate anthers. A pollen gun can be used to pollinate each female plant with the bulk of pollen. For bulking of pollen no special care is needed to avoid contamination because all pollen grains of the particular isolated field belong to the same population. For better control, pollinations can be performed alternately. For instance, in 1 day, pollination is in one direction (pollen
collected from field A) and in the next day pollination is in the opposite direction (pollen collected from field B).

Each female plant in each isolation field provides

(1) the inter-population half-sib family, obtained from the upper ear that was open pollinated by the opposite population, and
(2) the intra-population half-sib family, obtained by controlled pollination using a bulk of pollen from the same population (males in the other field).

Inter-population half-sib families (inter-HS) are evaluated in replicated yield trials and intra-population half-sib families (intra-HS) are used for recombination. Intra-HS from each population, corresponding to the selected inter-HS, are planted ear-to-row as females in one field. At the same time a bulk of seeds of the same families is planted as male rows in the other field. The second cycle starts by repeating the procedure of the first season. This technique provides simultaneously for recombination of selected progenies and for testcross material. Therefore, seeds obtained from the control hand-pollinated (lower) ear in the female rows in the first phase of the second cycle will correspond to the improved populations of the first cycle, A₁ and B₁ (see Table 12.10).

One cycle is completed every two seasons, which increases gain per year. Selection is also practiced in every generation, i.e., selection within female rows in 1 year and selection based on testcrosses in the other year. Prolificacy is under very strong selection intensity because when the second ear is protected its development is impaired in favor of the first ear. This reduces the number of plants with good seed set in the first cycles, but in subsequent cycles the success in obtaining good seed set on both ears increases.

Variations of the procedure are possible that can make it more applicable to the breeder’s facilities and available materials. For materials with a low level of prolificacy, both ear shoots could be protected and the second ear pollinated a few days before removing the bag of the first ear. Another variation may be to plant each population in an isolation block. The first ears are protected and the second ears are open pollinated with pollen from the same field (same population). The first ears are then pollinated with pollen collected from the other field. Such a modification may result in a greater number of plants with good seed sets but also with some disadvantages:

(1) Risk of contamination on the second ear is increased.
(2) Outcrossed ears will have a poorer seed set, limiting the number of replications in yield trials.

The method also can be modified to complete one cycle per season. In each year, outcrossed ears are evaluated in yield trials and two isolation blocks are planted for detasseling. In one field half-sibs obtained by hand pollination of A plants (intra-population half-sibs) are planted ear-to-row as females and a bulk of seeds from intra-population half-sibs (B) as males. In the other field population B is planted ear-to-row as females (half-sib family) with population A as male. When data of yield trials are available, only the best (selected) female rows are considered in each field. In
### Table 12.10 Description of the method of reciprocal recurrent selection based on half-sib progenies of prolific plants and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
</table>
| 1 (Winter) | **Testcross production in isolation**  
Population A is planted as females in a detasseling isolated block (Field 1), where population B is planted as male rows with a 1:2, 1:3, or 2:4 male:female ratio. The same is done in field 2 for population B  
Protect the second (lower) ear shoots of prolific plants in female rows in both fields. Pollen is bulked and used to pollinate the second ears in the opposite field (e.g., alternate pollination can be used)  
The first (upper) ears in each field are open pollinated  
Harvest both ears together with an appropriate identification. The upper ear is an inter-population half-sib family, whereas the lower ear is an intra-population half-sib family | — |
| 2 (Summer) | **Evaluate testcrosses**  
Conduct replicated trials of 100–200 testcrosses (half-sibs) from each field (A × B and B × A)  
Identify the best testcrosses from A × B and B × A | — |
| 3 (Winter) | **Recombine**  
Remnant seed of lower (hand pollinated) ears of population A harvested in season 1, Field 1, corresponding to the best A × B testcrosses, are planted ear-to-row as females in a detasseling block, as in season 1  
Male rows are a bulk of remnant half-sibs of population B that correspond to the selected testcross B × A  
Remnant half-sibs B related to the best B × A testcrosses are also planted ear-to-row in another isolation field where selected half-sibs A are bulked and planted as males | Include selected half-sib progenies in drought-managed environments and testcross nurseries for inbreeding and production of testcross seed |
| 4 (Summer) | Repeat as in season 1–3 | Include selected S₁ progenies in breeding and disease nurseries for selection and inbreeding |
the selected female rows both open-pollinated (inter-HS) and hand-pollinated (intra-HS) ears are harvested to initiate the next cycle (Paterniani and Vencovsky, 1978). Parental control is the same in one season/cycle and two seasons/cycle schemes.

12.5.4 Reciprocal Full-Sib Selection

Reciprocal full-sib selection was suggested by Hallauer and Eberhart (1970) as another method of inter-population improvement and for development of new lines from the two populations under selection. The operational procedures of reciprocal full-sib selection are similar to those proposed by Comstock et al. (1949) for RRS. The main difference is that full-sib progenies rather than half-sib progenies are evaluated. Reciprocal full-sib selection includes two breeding populations at a time for cyclical improvement. The bases of choosing breeding populations for reciprocal full-sib selection are the same as those for RRS. Because full-sib progenies are evaluated, a big advantage is that only one set of progenies is tested in reciprocal full-sib selection (in RRS two sets of half-sib progenies are evaluated). Hence, in reciprocal full-sib selection we can either sample twice as many plants in each of the two populations as in RRS or we can sample the same number of plants in each scheme and have half as many yield test plots. For example, if we sample 100 plants in each population in each selection scheme, we will have 200 half-sib progenies to test with RRS (100 for A tested with B and 100 for B tested with A) and 100 full-sib progenies for reciprocal full-sib selection. We will have equal numbers of yield test plots if we sample 200 plants from each population for reciprocal full-sib selection and 100 plants from each population for RRS. Therefore, in the first case we will have 50% fewer test plots with reciprocal full-sib selection and equal sampling of populations by both schemes. In the second case, we will have the same number of test plots for each scheme but twice as many plants are sampled with reciprocal full-sib selection.

Hallauer and Eberhart (1970) described reciprocal full-sib selection for populations including plants that produced seed on at least two ears (Hallauer, 1967a). Use of plants that produce seed on two ears will permit completion of one cycle of selection in 2 years (Table 12.11). Each cycle of selection is initiated in the summer season rather than in the winter season, as described for most of the other recurrent selection schemes. Unless the off-season nursery is in another temperate zone, the quantity of seed set on the second ear is usually not satisfactory. If the off-season nursery can be used to produce S1 progenies and full-sib crosses each cycle of selection can be completed in 2 years, but an additional cycle of recombination can be completed between each cycle of selection.

If we assume the summer season is used to produce S1 and full-sib progenies, the two base populations are planted in alternate rows in the breeding nursery. Plant density used will depend on the frequency of two-ear development in the breeding nurseries. Lower plant densities will help in the expression of prolificacy (Carena et al., 1998). In the Iowa program 40,000–60,000 plants/ha have been used for the two populations undergoing reciprocal full-sib selection. Frequency of two-ear development was consistently high in both populations (BS10 and BS11), but it
Table 12.11 Description of the use of different seasons for conducting reciprocal full-sib recurrent selection with prolific plants in temperate zones and integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
</table>
| 1 (Summer) | *Produce S₁ progenies and full-sib crosses*  
Plant the two populations (C₀) under selection in alternate rows in the breeding nursery  
Self-pollinate one ear and cross the other ear  
Practice selection among S₀ plants at pollination and harvest | — |
| 2 (Summer) | *Evaluate full-sib crosses*  
Grow the full-sib crosses in replicated yield trials at 3–4 locations  
Select (e.g., index selection for multiple traits) 20–30 top full-sib progenies for recombination | *Grow pairs of S₁ progenies in breeding nurseries to initiate line and hybrid development* |
| 3 (Winter) | *Recombine*  
Recombine the selected S₁ progenies of the selected full-sib crosses (e.g., 20–30) to form C₁ Syn 1 for each of the two populations. Use the bulk-entry method of recombination for each population | *Include selected S₂ crosses in winter nurseries for inbreeding, drought screening, and testcrossing* |
| 4 (Summer) | *Produce S₁ progenies and full-sib crosses*  
Plant the C₁ Syn 1 populations in alternate rows in the breeding nursery  
Self-pollinate one ear and cross the other ear  
Practice selection among S₀ plants at pollination and harvest | *Grow S₂ or S₃ (if winter nursery available) progenies of selected full-sib crosses in breeding, disease, and testcross nurseries for additional selection and inbreeding* |
| 5 (Summer) | Repeat as in seasons 2 and 3 | *Grow selected S₃–₅ progenies in breeding nurseries and replicated testcross trials* |

*The other option is to only self-pollinate and select among S₁ progenies for both populations in order to produce crosses between pairs of S₁ progenies the following season. A disadvantage is one less season. Among the advantages are larger seed production for testing across several environments and the possibility of obtaining S₃ lines at the time of data collection (e.g., non-prolific populations).*

will depend to some extent on environmental conditions, such as moisture and temperature. Selected S₀ plants of the two populations, planted in alternate rows, are chosen at pollination. The most important trait usually is the time of flowering. The S₀ plants that have the same flowering dates are prepared (by covering silks on both ear shoots) for pollination with the first pollination made on the second ears. If the cross-polinations are made first, the tassels are covered and used to pollinate the second ears the next day. The same day the second ears are pollinated the tassels are covered again and the top ears are self-pollinated the second day.
It has been found that if the second ears are pollinated first (rather than the top ears first), more successful seed sets are obtained. It is necessary to accurately label the bags of the two plants used to make the cross-pollinations in order to bulk the two ears included in producing the full-sib progenies, similar to identifying full-sib crosses in intra-population recurrent selection. The primary reason the second ears were used to produce the full-sib progenies was to increase the frequency that selfed seed was available for each plant included in the crosses. If one of the second ears is lost, we usually have sufficient seed on one of the second ears to conduct a two-replicate yield trial at three locations (about 300–500 seeds). Saving seed from only plants that produce seed on both ears exerts strong selection pressure for prolificacy. If the plants do not have seed on two ears, they are discarded. Continued selection fixes a greater frequency of two-ear plants. Seed from S\textsubscript{1} and full-sib progenies are harvested. In some cases plant pairs have weak stalk and should be discarded. All bags are labeled and one can check to determine if the particular pairs of S\textsubscript{0} plant crosses were successful. Full-sib ear pairs are bulked (or in some instances seed from only one ear) for yield trials in season 2 (Table 12.11). Each pair of S\textsubscript{1} ears is identified for each cross so that we harvest four ears for each pair of S\textsubscript{0} plant crosses. The full-sib progenies, say 150–200, are grown in replicated yield trials of two replications at three or four locations. Data of the full-sib progenies are summarized for all locations and the highest yielding full-sib progenies are identified. Or, alternatively, an index selection for multiple traits could be used. Following the ISU program in the 1990s, the NDSU maize breeding program often utilizes the heritability index based on yield, grain moisture at harvest, stalk lodging resistance, and test weight for reciprocal full-sib progeny evaluation but often at plant densities closer to 85,000 – 100,000 plants/ha as enough seed is produced in pairs of S\textsubscript{1} progenies. However, the index changes depending on the target state region or if a particular set of traits is targeted for improvement (e.g., extractable starch and other grain quality traits). The S\textsubscript{1} seed of the full-sib progenies is planted in the winter nursery for recombination. As for RRS, there will be two sets of S\textsubscript{1} progenies for recombination, one for each of the two populations undergoing reciprocal full-sib selection. After recombination, the next cycle of reciprocal full-sib selection is initiated in the C\textsubscript{1} Synthetic 1 generation, which is season 4 (Table 12.11). The same season that another cycle of reciprocal full-sib selection is initiated, the S\textsubscript{1} progenies included for recombination can be included in the breeding and testcross nurseries to initiate the inbred line development process. The S\textsubscript{1} progenies have not had any previous selection except for combining ability in the full-sib progeny tests. Choice of testers and stage of testcrossing would be the same as described for RRS.

Reciprocal full-sib selection was designed to maximize selection for specific combining ability in the development of single-cross hybrids (Hallauer, 1967a, b). Instead of testing for general combining ability and then searching for a specific single-cross combination, emphasis was given to selecting for specific combining ability in each generation of inbreeding. Selected progenies were not included in the testcross nursery but only in breeding and pest nurseries. The S\textsubscript{1} progenies of selected full-sib progenies are planted ear-to-row in pairs that correspond to the respective S\textsubscript{0} plants included in full-sib progenies tested in season 2 (Table 12.11). Each pair of S\textsubscript{1} progenies is planted in 16–25 plant rows in the breeding nursery.
Within each pair of progeny rows, three to five pairs of S₁ plants are selfed and crossed as described for season 1. Some of the pairs are discarded before pollination for certain agronomic traits and first brood European corn borer. All plants pollinated in the S₁ generation are inoculated for stalk rot organisms. Seed is harvested only from plants that have good stalks and produce seed on both ears. One ear is S₂ seed and the second ear is the full-sib seed produced between S₁ plants. After full-sib progeny yield trial data are available, many progenies are discarded but only those that exhibit good yield potential as full-sib progenies are continued for additional inbreeding, selection, and testing. In some instances a pair of S₁ progenies may have been discarded for some trait but had high yield in the full-sib progeny yield tests. If the S₁ progenies were marginal for some traits but high yielding, they may be included for recombination but not in the breeding nursery in the S₂ generation. Machine harvesting of full-sib progeny trials will eliminate progenies with a high frequency of stalk breakage and usually confirm the data on poor stalk quality of S₁ progenies in the breeding nursery. Opportunities exist for selecting for good stalk quality in the breeding nursery and the full-sib progeny trials. Inbred development phase of reciprocal full-sib selection is continued as described by Hallauer (1973b), Fig. 12.9.

Fig. 12.9 Integration of selection methods in two populations for contributing to long-term (stratified mass selection for adaptation), mid-term (reciprocal full-sib recurrent selection to maximize genetic improvement), and short-term (pedigree selection for line development and hybrid identification) breeding goals.

Although specific combining ability was emphasized with the use of reciprocal full-sib selection, Hoegemeyer and Hallauer (1976) found that elite inbred lines developed by reciprocal full-sib selection also had good general combining ability with other elite lines. There was concern that new inbred lines developed by testing only particular pairs of lines may have limited usefulness because they may not be
acceptable for use as either male or female parent seed stock in production of the single-cross hybrid. Hoegemeyer and Hallauer used the design II mating scheme to test single-cross hybrids of inbred lines tested through the \( S_4 \) generation via reciprocal full-sib selection. Twenty-four pairs of selected lines in the \( S_7 \) generation were included in 96 single crosses. Primary interest was to determine how the diagonal single crosses of the design II mating scheme compared with the off-diagonal single crosses. The diagonal single crosses were those that had been tested and selected by reciprocal full-sib crosses, whereas the off-diagonal single crosses (not previously tested) were produced among the selected lines. The mean yield of the diagonal crosses (92 q/ha) for the six sets was significantly greater than the mean yield of the off-diagonal crosses (88 q/ha). Single-crosses selected by reciprocal full-sib selection, therefore, were on the average greater yielding than the previously untested off-diagonal crosses. But in each set at least one of the off-diagonal crosses was greater yielding than the diagonal single crosses. Analysis of variance of the design II mating scheme permits the estimation of general and specific combining ability effects and tests of mean squares for general and specific combining ability sources of variation. Hoegemeyer and Hallauer (1976) found that specific combining ability mean squares were significant in four of the six sets of single crosses. General combining ability mean squares were significant in all instances. Estimates of specific combining ability effects for the 24 selected diagonal single crosses, however, were predominantly positive and significant (16 of 24); whereas estimates of specific combining effects for the 72 off-diagonal untested single crosses were significantly positive in 21 instances, significantly negative in 38 instances, and not significantly different from zero in 13 instances. It seems that reciprocal full-sib selection was effective in selecting for specific combining ability, but that specific effects were relatively small compared with general effects. Hence elite lines developed by reciprocal full-sib selection that emphasized selection for specific combining ability also had good general combining ability with other elite lines. Concern that inbred lines developed by reciprocal full-sib selection would combine well only with the particular line with which it had been tested was, therefore, not valid. Conclusions were that although reciprocal full-sib selection could effectively select for specific combining ability (non-additive effects), the procedure also was effective for general combining ability (primarily additive effects) and general combining ability was generally more important than specific combining ability.

Techniques of reciprocal full-sib selection were initially described with the use of two-eared plants (Hallauer, 1967a). Reciprocal full-sib selection also is usable if either two-eared germplasm is not available or the maize breeder does not want to use a breeding procedure that involves multiple pollinations on the same plant (Hallauer, 1967b, 1973b; Marquez-Sanchez, 1982). The chronological description of reciprocal full-sib selection with use of primarily one-eared maize plants is given in Table 12.12. Use of one-eared rather than two-eared plants increases cycle time from 2 to 3 years. Effect of increased cycle time is compensated for by the additional selection that can be practiced for agronomic and pest traits. In addition, the amount of seed that can be produced for multi-location trials is large enough to significantly reduce genotype by environment interactions. Pairs of \( S_1 \) progeny rows have several plants bearing ears to produce full-sib seed. Theoretically, increasing cycle time
Table 12.12 Description of the use of different seasons for conducting reciprocal full-sib recurrent selection without prolific plants in temperate zones and the integration with inbred line development

<table>
<thead>
<tr>
<th>Season</th>
<th>Population improvement program</th>
<th>Cultivar development program</th>
</tr>
</thead>
</table>
| 1 (Winter) | *Self-pollination*  
Self-pollinate $S_0$ plants in the two populations (C0) under selection. Add date of pollination on bags  
Practice selection among $S_0$ plants at pollination and harvest | — |
| 2 (Summer) | *Produce full-sib crosses*  
Plant the $S_1$ progenies of the two C0 populations in alternate rows in the breeding and pest nurseries according to the date of pollination of $S_0$ plants  
Produce full-sib progenies on a portion of the $S_1$ plants and self-pollinate selected $S_1$ plants (see discussion) | — |
| 3 (Summer) | *Evaluate full-sib crosses*  
Grow the full-sib crosses in replicated yield trials at 6–8 locations  
Select (e.g., index selection for multiple traits) 20–30 top full-sib crosses  
$S_2$ progenies can be included in the breeding nurseries for additional self and cross-pollination | |
| 4 (Winter) | *Recombination*  
Recombine remnant seed of 20–30 $S_1$ or $S_2$ progenies based on data of the selected full-sib crosses to form C1 Syn 1 cycle. Use the bulk-entry method of recombination for each population  
Include selected $S_3$ progenies in winter nurseries for inbreeding, drought screening, and testcrossing | |
| 5 (Summer) | *Recombination*  
Plant the C1 Syn 1 populations to produce the C1 Syn 2 cycle populations  
To reduce the number of pollinations for recombination, after each pollinating, remove shootbags and detassel males and females  
Grow $S_3$ or $S_4$ (if winter nursery available) progenies of the selected full-sib crosses in the breeding, disease, and testcross nurseries for additional selection | |
| 6 (Winter) | *Self-pollination*  
Self-pollinate $S_0$ plants in the two C1 Syn 2 cycle populations. Add date of pollination on bags  
Practice selection among $S_0$ plants at pollination and harvest  
Grow selected $S_{3-5}$ progenies in breeding nurseries and replicated testcross trials | |
| 7 (Summer) | Continue as in seasons 2–3  
Grow selected $S_{3-5}$ progenies in breeding nurseries and replicated testcross trials to determine combining ability | |
from 2 to 3 years will decrease predicted gain on a per-year basis, but additional selection will result in improved material for inbred line development.

Reciprocal full-sib selection with use of one-eared plants is initiated in the winter nursery by selfing selected S₀ plants in the two broad genetic base populations. Date of pollination should be recorded on the pollination bag. At harvest, selected ears with their dates of pollination are saved and prepared to plant ear-to-row in the summer nursery (season 2, Table 12.12). The S₁ progenies will need to be planted in paired ear rows: one for population A paired with one for population B. The purpose for recording dates of pollination of the S₀ plants is to pair the S₁ progenies that have similar dates of flowering. If dates of pollination are not recorded, split plantings for each row could be made to ensure timing of the tassels and silks at pollination. Paired S₁ rows also should be planted in additional nurseries for screening after flowering (e.g., to eliminate those that are obviously susceptible to the primary pests, tillers, grain filling, lodging, drought) and before flowering (e.g., first brood European corn borer, seedling vigor, emergence).

Because only one-eared plants are available, the same plants cannot be used to produce the self (S₂) and full-sib seed and different S₁ plants within each pair of S₁ progenies will be self-pollinated and cross-pollinated. Thus there will be some sampling differences of genotypes used for the self- and cross-pollinations. But the sampling should not be too great because the genotypic array of the S₁ progeny represents the genotype of the S₀ plant that was self-pollinated. Self-pollinate selected plants in each row. Similarly, cross-pollinate other selected plants between each pair of S₁ progeny rows. Or (preferably) self-pollinated plants can be used to cross-pollinate other selected plants. All pollinations can be made the same day. It will minimize errors if one portion of the S₁ progeny rows is self-pollinated and the other portion is cross-pollinated. Clearly mark pollinations to indicate self- or cross-pollination. Another option would be to use plain pollination bags for self-pollinations and striped color bags for cross-pollinations. After pollination, inoculate all pollinated plants within each row with diseases, preferably in a duplicated disease nursery. Harvest self- and cross-pollinated ears in pairs of S₁ progenies that have acceptable stalks, ears, kernel type, etc. Self-pollinated ears need labeling for each progeny row. Harvest cross-pollinated ears and shell in bulk in preparation for yield trials. Again, one great advantage of this procedure with one-eared plants is that adequate pollinations can be made to ensure adequate seed for the yield trials. If three to five plants are cross-pollinated for each plant self-pollinated, adequate quantities of seed should be available for yield testing and the differences due to sampling should not be serious. The full-sib progenies of S₁ progenies are tested in season 3 (Table 12.12).

Other procedures of reciprocal full-sib selection with use of one-eared plants are the same as those described with use of two-eared plants. The primary difference is the self-pollination in the first season to initiate each cycle of recurrent selection. One additional season is required but effective selection can be made among S₁ progenies, whereas only selection among S₀ plants was possible with use of two-eared plants. Labeling is reduced with use of one-eared plants. Also, fewer opportunities for errors exist at pollination, harvest, and shelling. If the winter nursery is used
to initiate the next cycle of selection, another recombination can be made in the summer season to produce the Synthetic 2 generation. Reciprocal full-sib selection with use of one-eared plants may be more amenable to more applied breeding programs than with use of two-eared plants (Marquez-Sanchez, 1982).

Hallauer (1978) and Obilana et al. (1979b) have reported on the initial effectiveness of reciprocal full-sib selection in Iowa Two-Ear Synthetic (BS10) and Pioneer Two-Ear Composite (BS11). Selection was effective in significantly improving yield of the populations themselves and the cross of the two populations (Table 12.13). Observed progress in the population cross was in close agreement with predicted yield, which was based on components of variance estimated for the inter-population cross of the C0 populations. Progress for stalk and root lodging and grain moisture at harvest also was encouraging. Although differences between the original and selected cycles were not all significant, the changes were in the desired direction for just the first three cycles of selection. Stalk lodging was decreasing and root lodging and grain moisture at harvest also tended to decrease, but not significantly in all instances. Selection was effective in changing populations for the primary trait, yield, and correlated agronomic traits that are important in applied breeding programs. Selection, therefore, continued.

Frequency distributions for grain yield and grain moisture of full-sib families produced and tested at the C0, C7, and C13 cycles of selection are included in Figs. 12.10 and 12.11 to illustrate response to reciprocal full-sib selections in BS10 and BS11. There were 144 full-sib families tested in 1964 for the C0 cycle, 145 full-sib families tested in 1984 for the C7 cycle, and 214 full-sib families tested in 1999 for C13 cycle. The same six hybrid checks (three single and three double crosses) were included in all trials for all cycles of selection to provide a consistent point of reference and distributions are relative to the mean of the six check hybrids. Distributions for grain yield for C0 and C13 cycles of selection were essentially transposed relative to the mean of the six check hybrids (Fig. 12.10). The theoretical effects of selection illustrated in Fig. 12.3 agree with the data shown in Fig. 12.10. As stated early in the chapter only two of the 144 C0 × C0 full-sib families exceeded the mean of the six check hybrids, whereas only 17 of the 214 C13 × C13 full-sib families had yields less than the mean of the six hybrid checks. Grain yield was the more important trait considered in selection but proper maturity and good root and stalk strength are required in modern maize production. Although greater grain yields are often associated with later maturity, the data illustrated in Fig. 12.11 show that average grain moisture and variability have been reduced from C0 to C13 cycles of selection.

Cycle 18 will be completed in 2010 and inbred lines (B77, B79, B98, B113, and B115) have been released after improved selections of these improved populations were incorporated in pedigree selection programs. The trends shown in Table 12.13 after three cycles of selection continued but at a slower rate.

Good stalk quality is essential if populations are to be useful as breeding material and if derived inbred lines are to be useful in hybrids. Table 12.13 shows that stalk lodging decreased in the populations themselves and in the population cross. Figure 12.12 shows progress has been made for improved stalk quality of full-sib progenies.
Table 12.13  Observed comparisons of BS10 and BS11 maize populations and crosses for four traits after the initial three cycles of reciprocal full-sib recurrent selection (Obilana et al. 1979a, b)

<table>
<thead>
<tr>
<th>Trait</th>
<th>BS10</th>
<th>BS11</th>
<th>BS10 × BS11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C0</td>
<td>C3</td>
<td>d</td>
</tr>
<tr>
<td>Yield, q/ha</td>
<td>53.9</td>
<td>62.8</td>
<td>8.9 ± 3.6</td>
</tr>
<tr>
<td>Grain moisture, %</td>
<td>22.1</td>
<td>19.9</td>
<td>-2.2 ± 1.0</td>
</tr>
<tr>
<td>Root lodging, %</td>
<td>5.6</td>
<td>8.3</td>
<td>2.7 ± 3.6</td>
</tr>
<tr>
<td>Stalk lodging, %</td>
<td>18.1</td>
<td>12.8</td>
<td>-5.3 ± 3.3</td>
</tr>
</tbody>
</table>
Fig. 12.10  Frequency distributions for grain yield of full-sib families for the original BS10 and BS11 cycle (C0 × C0, left line) and after seven (C7 × C7, dotted line) and 13 cycles (C13 × C13, right line) of reciprocal full-sib selection relative to the mean of six hybrids included as checks for all cycles of selection

Fig. 12.11  Frequency distributions for grain moisture of full-sib progenies for the original BS10 and BS11 cycle (C0 × C0, right line) and after seven (C7 × C7, dotted line) and 13 (C13 × C13, left line) cycles of reciprocal full-sib selection relative to the mean of six check hybrids
The means of C0 and C7 full-sib progenies exceeded the mean of the six check hybrids. However, change in the distributions is more evident. Only 6.9% of C0 full-sib progenies had less stalk lodging than the mean of the six check hybrids, but 20.6% of C7 full-sib progenies had less stalk lodging than the mean of the hybrid checks. The mean of the 144 C0 full-sib progenies was nearly four standard errors above the mean of the checks, whereas the mean of the 145 C7 full-sibs was nearly one standard error higher. Selection and testing in each generation of inbreeding provide good opportunities for developing improved stalk quality by

1. inoculation and selection among and within progenies in the breeding nursery, and
2. taking notes and measuring harvestable yield from mechanically harvested yield trials.

In the breeding nursery information is obtained on progenies themselves and in yield trials information is obtained on the combining ability for stalk quality in full-sib progeny crosses. Use of plants that produce seed on two ears puts additional stress on plants for good stalk quality. It seems that the physiological limits of increased kernels and improved stalk quality can be combined. In fact, the trend continued with root (−11.8%) and stalk (−8.7%) lodging which also were reduced after 13 cycles of reciprocal full-sib family selection.

Reciprocal full-sib selection (like other recurrent selection schemes) can be used to meet short-, intermediate-, and long-term objectives of a breeding program.
Mass selection in BS10 and BS11 has been conducted in Iowa to provide backup populations for the reciprocal full-sib program (Weyhrich et al., 1998a). Mass selection was conducted in isolated plantings of about 0.4 ha and ears are selected from plants that produce two ears on standing plants. If in the future it seems that genetic variation is being reduced by reciprocal full-sib selection, mass selected populations can be crossed with populations under reciprocal full-sib selection. The mass selection portion of the program satisfies long-term objectives. Long-term objectives are improvement by mass selection, intermediate-term objectives are improvement of populations by reciprocal full-sib selection, and short-term objectives are line development from each cycle of reciprocal full-sib selection.

Empirical data comparing RRS and full-sib reciprocal selection are limited to those shown in Hallauer and Carena (2009), and progress reports by Keeratinijakal and Lamkey (1993) for half-sib RRS in BSSS and BSCB1 and by Eyherabide and Hallauer (1991) for full-sib RRS in BS10 and BS11. Realized response for eight cycles of reciprocal full-sib selection in BS10 and BS11 was 7.5% per cycle, which compares favorably with the response of 7 and 6% reported by Darrah et al. (1978) and Gevers (1975) for three cycles of RRS. Some theoretical guidelines were provided by Jones et al. (1971) for response to selection expected by RRS and full-sib recurrent selection. For the situation of equal additive and dominance variances, the phenotypic variation among half-sib families was one-third the phenotypic variation among full-sib families. To compensate for the lower phenotypic variation, the selection differential would need to be 1.7 times greater with full-sib than with half-sib RRS. Comparisons with actual estimates of components of variance showed that the selection differential should be 1.2 times greater for reciprocal full-sib selection. Inter-population estimates of additive genetic and dominance components of variance and their interactions with environments were reported by Obilana et al. (1979a) for BS10 and BS11. Their estimates were used to calculate the phenotypic variance for half-sib and full-sib progenies. For the situation of two replications in each of three environments the full-sib phenotypic standard deviation was 1.24 times greater than the half-sib phenotypic standard deviation. Jones et al. (1971) concluded that reciprocal full-sib selection was favored at lower selection intensities and for large environmental variance relative to total genetic variance but declined when selection intensity increased.

12.5.5 RRS with Use of Inbred Lines

Because it seems that specific combining ability is of minor importance, Russell and Eberhart (1975) and Walejko and Russell (1977) suggested that RRS be conducted with use of inbred lines as testers. Choice of inbred lines to use as testers should be the same as choice of two populations to include for RRS. Hence, tester lines should conform to the heterotic pattern expressed in hybrids. Description of RRS with use of inbred testers is the same as shown for RRS based on testcrosses of half-sib families (Section 12.5.2) except inbred lines rather than opposing populations are used as respective testers. Instead of crossing plants from population A with
plants from population B (the tester), plants from population A are crossed with an inbred tester extracted from population B. The reverse situation would be used for population B. If the two populations are of Reid and Lancaster Surecrop origin, the Reid population would be crossed with an inbred line tester (e.g., Mo17 or Oh43) of Lancaster Surecrop derivation and the Lancaster Surecrop population would be crossed with an inbred line tester (e.g., B73 or A632) of Reid derivation. The same conditions hold if flint and dent populations are undergoing selection. Again, current industry testers would be a good choice.

After choices of populations and inbred line testers are decided, mechanics of the breeding operation follow the description given in Table 12.9. We would self selected S₀ plants in both populations in the off-season (winter nursery) in season 1. Selection among and within S₁ progenies of each population would be performed in season 2. Then, cross selected S₁ plants to the inbred tester. Also self each S₁ plant used to make the testcrosses. Grow the two sets of testcrosses in replicated yield trials. The S₂ progenies of S₁ plants used to make the testcrosses can also be included in the breeding nursery for another generation of selection and inbreeding. By the time the testcross data are available, S₃ generation progenies are available for the section of the breeding nursery focused on inbred line development. On the basis of testcross trials, select S₁ or S₂ progenies for recombination to form the next cycle. The choice of progenies (S₁ or S₂) to recombine is the same as discussed for RRS. Another generation of recombination can be made in the summer season before initiating selfing of S₀ plants in the recombinant population. Each cycle of RRS with use of inbred testers would require a minimum of 3 years per cycle, which includes two generations of recombination (seasons 4 and 5).

Use of inbred line testers in RRS provides S₂ lines for the breeding nursery that have had selection among and within S₁ progenies and combining ability tests with the inbred line tester. Choice of tester is important. Use of inbred parental lines of a highly heterotic hybrid of a particular region (e.g., B73 × Mo17, A619 × A632, TR3622GT × TR3030CBLL, ND400 × TR3026, ND2000 × TR1017Bt, etc.) seems appropriate. Although inbred line testers were suggested by Walejko and Russell (1977), single crosses of the appropriate pedigree also can be used to reduce the concern of having adequate seed for the yield trials. Because of the seemingly small importance of specific combining ability, testers can be changed to meet the current use of specific lines in hybrids. A particular line or lines may be suitable for two or three cycles of selection and then changed if the line is not useful or being used in hybrids. Since general combining ability seems of greater importance than specific combining ability (e.g., Horner et al., 1973, 1976; Hoegemeyer and Hallauer, 1976) even with use of a specific inbred line, the change in lines used as testers should not affect continued selection for improvement of populations.

RRS with use of an inbred line was designed for applied breeding programs (Russell and Eberhart, 1975). Accumulated information of relative importance of specific vs. general combining ability with different types of testers was instrumental in developing the recurrent selection scheme. Elite lines are required as parental stocks in production of single-cross hybrids; consequently, an important objective of most applied breeding programs is the improvement of lines for one
or both sides of the hybrid pedigree. Most pedigree selection programs emphasize improvement of one or both lines of the pedigree. RRS by use of an inbred line tester is a useful recurrent selection scheme for supplementing pedigree selection programs right away. Advantages and disadvantages for use of an inbred line as tester for recurrent selection are discussed in Chapter 8. Comparisons of advantages and disadvantages of the method relative to RRS were discussed by Comstock (1979).

### 12.6 Additional Considerations for Germplasm Improvement

Choice of population(s) to include for improvement and choice of recurrent selection scheme(s) to use for the improvement of the population(s) are decisions that must be made before the improvement program is initiated. Populations to include for improvement should have adequate genetic variability, have high mean, and manifest heterosis in crosses, particularly if some form of inter-population recurrent selection is proposed. Information may not be explicitly available for populations available, but a combination of experimental data, personal experience, and the experience of others may provide leads to the most potentially useful populations to include for improvement. It seems that adequate genetic variability is present in most populations to expect progress from selection, but some populations may have more useful genetic variability than others. For instance, Iowa Stiff Stalk Synthetic seems to be a more useful population for developing new inbred lines than Iowa Corn Borer Synthetic No. 1, although the populations seem to have similar genetic variability. A population that has high mean yield performance, other traits being equivalent, would be preferable to one that has low mean yield. Equivalent rates of improvement may be made for both low and high mean yield populations, but the low mean yield population would not be equivalent to the high mean yield population. Even if the rate of improvement was somewhat greater for the low mean yield population, it probably would not yield as many potentially useful lines as the higher mean yield population. If the breeding program is not committed to use of inbred lines in hybrids, use of populations that have high mean performance also would be more feasible. It is important to capitalize on the expression of heterosis either in crosses of inbred lines extracted from populations undergoing recurrent selection or in population hybrids for breeding programs not committed to use of inbred lines. Manifestation of heterosis should be considered for both intra- and inter-population improvement programs. If two populations are undergoing recurrent selection, it is important that heterosis is expressed in their cross. Chapter 5 emphasizes importance of selection among populations to initiate a recurrent selection program and formulas are given in Chapter 10 that provide the basis for prediction of composite variety means or the mean of a cross between two composites, using data from variety diallel crosses. Use of prediction procedures permits selection not only among existing varieties but also among populations that would result from crosses involving such a set of varieties.
Choice of recurrent selection scheme to use for improvement of populations depends on the breeder’s objectives. It seems that all recurrent selection schemes effectively improve populations chosen for improvement. Based on realized gain per cycle, all recurrent selection schemes are nearly equally effective. Summaries of selection results for those schemes included in Chapter 7 were given by Sprague and Eberhart (1977), Hallauer et al. (1988), Hallauer (1992), and Hallauer and Carena (2009). Rate of realized gain is about 3–7% per cycle for intra- and inter-population recurrent selection schemes. Different populations were included in different recurrent selection schemes conducted in different areas, but realized gain per cycle on the average is not greatly different. Rate of gain per year would be different because different numbers of years are required for different recurrent selection schemes. Comparisons for different methods of recurrent selection listed in Table 12.14 are averages for recurrent selection in different populations by the 1980s which are similar to those reviewed lately.

**Table 12.14** Summary of the average genetic gain per cycle for several recurrent selection schemes

<table>
<thead>
<tr>
<th>Selection scheme</th>
<th>Chapter 7</th>
<th>Sprague and Eberhart (1977)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intra-population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>2.9</td>
<td>3.4</td>
</tr>
<tr>
<td>Ear-to-row</td>
<td>6.5</td>
<td>3.8</td>
</tr>
<tr>
<td>Full-sib</td>
<td>3.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Testcross</td>
<td>5.4</td>
<td>2.8</td>
</tr>
<tr>
<td>S1</td>
<td>6.4</td>
<td>4.6</td>
</tr>
<tr>
<td>S2</td>
<td>—</td>
<td>2.0</td>
</tr>
<tr>
<td>Average</td>
<td>5.0</td>
<td>3.3</td>
</tr>
<tr>
<td><strong>Inter-population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reciprocal</td>
<td>4.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Darrah et al. (1978)a</td>
<td>7.0</td>
<td>—</td>
</tr>
<tr>
<td>Gevers (1975)a</td>
<td>6.0</td>
<td>—</td>
</tr>
<tr>
<td>Paterniani &amp; Vencovsky (1978)a</td>
<td>3.5</td>
<td>—</td>
</tr>
<tr>
<td>Testcross</td>
<td>4.2</td>
<td>3.5</td>
</tr>
<tr>
<td>Reciprocal full-siba</td>
<td>5.0</td>
<td>—</td>
</tr>
<tr>
<td>Average</td>
<td>5.3</td>
<td>3.2</td>
</tr>
</tbody>
</table>

a Not included by Sprague and Eberhart (1977)

Darrah et al. (1978) reported on a comprehensive series of selection studies conducted for three populations, Kitale Composite, Ec573, and KII (Eberhart et al., 1967). Because different methods of recurrent selection were conducted in the same populations, direct comparisons could be made for realized gain (Table 12.15). Darrah et al. (1978) presented their results on a per-year basis, but the data were converted to gain per cycle to make them equivalent to the average realized gains given in Table 12.14. Except for RRS, realized gains per cycle were similar to average realized gains given for different methods of recurrent selection. Realized gain
Table 12.15 Response to selection for yield in a breeding methods study in Kenya, East Africa (Darrah et al., 1978)

<table>
<thead>
<tr>
<th>Selection methods</th>
<th>Gain per cycle (q/ha)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCA-M9C6</td>
<td>0.38</td>
<td>0.8</td>
</tr>
<tr>
<td>M10C6</td>
<td>0.93**</td>
<td>1.6</td>
</tr>
<tr>
<td>M17C6</td>
<td>−0.70</td>
<td>−</td>
</tr>
<tr>
<td>Ear-to-row</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KII</td>
<td>0.83**</td>
<td>1.8</td>
</tr>
<tr>
<td>Ec573</td>
<td>2.59**</td>
<td>6.6</td>
</tr>
<tr>
<td>KCA-E3C6</td>
<td>0.98**</td>
<td>2.1</td>
</tr>
<tr>
<td>E4C6</td>
<td>0.82**</td>
<td>1.4</td>
</tr>
<tr>
<td>E5C6</td>
<td>1.04**</td>
<td>2.2</td>
</tr>
<tr>
<td>E6C6</td>
<td>0.56**</td>
<td>1.2</td>
</tr>
<tr>
<td>E7C6</td>
<td>1.40**</td>
<td>2.5</td>
</tr>
<tr>
<td>Half-sib(^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCA-H14C3</td>
<td>1.26</td>
<td>2.0</td>
</tr>
<tr>
<td>HL15C3</td>
<td>0.62</td>
<td>1.0</td>
</tr>
<tr>
<td>HH16C3</td>
<td>−3.26**</td>
<td>−</td>
</tr>
<tr>
<td>S(_1)(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCAC3</td>
<td>−0.50</td>
<td>−</td>
</tr>
<tr>
<td>Full-sib</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCAC3</td>
<td>1.60*</td>
<td>2.4</td>
</tr>
<tr>
<td>Reciprocal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KII3</td>
<td>−0.04</td>
<td>−</td>
</tr>
<tr>
<td>Ec573C3</td>
<td>1.94**</td>
<td>2.5</td>
</tr>
<tr>
<td>(KII × Ec573)C3</td>
<td>4.18**</td>
<td>7.0</td>
</tr>
</tbody>
</table>

* and ** indicate significance at \( P = 0.05 \) and \( 0.01 \), respectively

\(^a\) Population by tester crosses and not populations themselves

\(^b\) Based on bulk S\(_1\) lines rather than random mated populations

for RRS was 7%, which was similar to the programs in the USA but also similar to that reported by Gevers (1975) for Teko Yellow and Natal Yellow Horsetooth in South Africa (6%). Weyhrich et al. (1998a) compared the responses for seven recurrent selection methods in the BS11 maize population, ranging from mass selection for increased prolificacy, inbred (S\(_1\) and S\(_2\)) progeny selection, half-sib and use of inbred tester, full-sib, and reciprocal full-sib selection. Genetic response was realized by all selection methods, but the greatest response for grain yield was 4.5% per cycle for 5 cycles of S\(_1\) progeny selection and the lowest for 10 cycles of mass selection (0.6% per cycle). Hence the choice of recurrent selection scheme depends on how the particular scheme complements other phases of the breeding program.

Initial selection among and within progenies for agronomic traits may be preferable before obtaining a measure of combining ability for either non-inbred or S\(_1\) progenies; for this instance, S\(_1\) or S\(_2\) recurrent selection may be preferable to some type of half-sib recurrent selection. Choice of intra- or inter-population recurrent selection for two populations also depends on the breeder’s preference. The S\(_1\) or S\(_2\) recurrent selection would not be appropriate for generally accepted schemes of inter-population recurrent selection but it could be used for two populations that meet the same criteria. The logical sequel would be either to cross improved populations to produce the population cross or to extract lines from improved populations...
for use in hybrids regardless of whether intra- or inter-population recurrent selection was used.

Decisions of choice of populations and recurrent selection schemes are not always obvious, but adjustments can be made during the course of the program if the breeder feels the wrong choices were made initially. Changes in a recurrent selection scheme can be made from say half-sib recurrent selection to S1 recurrent selection in subsequent cycles if the breeder feels that the scheme currently used is not meeting objectives. For example, a broad genetic base population can be effectively improved through half-sib family selection if the objective is to use the improved population directly for grain production. However, if the population was never submitted to a strong inbreeding pressure it would not have immediate value as a source of vigorous inbred lines, since greater inbreeding depression is expected to occur after selfing. Should the breeder change objectives during the program and want to use the improved population as a source of inbred lines for hybrids, one or more cycles of S1 family selection would provide an opportunity for selection against deleterious recessive genes and make the population more suitable for extraction of inbred lines. If the base population originated from inbred material, such as synthetics from inbred lines, use of inbred family selection would have less pronounced results than if it had never been submitted to strong inbreeding. Additional germplasm can be included in populations under recurrent selection to rectify some obvious weaknesses. Populations could be kept open and not closed. If wise choices were made for populations to include to initiate recurrent selection, the loss of future genetic gains from changes and/or adjustments in selection methods would be less than to replace the populations under selection. Past gains would be lost if one started selection anew in different germplasm sources. Proper and wise choices of germplasm are essential when one initiates long-term selection programs. One needs to remain flexible and adapt to changes when conducting recurrent selection programs, but frequent changes in germplasm would negate past genetic gains.

Utilizing germplasm and recurrent selection for basic studies is encouraged. Suggestions for modifying populations and recurrent selection schemes would not be detrimental if one is interested primarily in obtaining basic information on breeding methods, the same way QTL studies have mostly worked with elite × poor combinations. In fact, recurrent selection methods would also be essential to demonstrate current hypotheses based on theories and simulations (e.g., genome-wide selection).

In this book, however, we are primarily interested in using recurrent selection schemes as adjuncts and/or essential components to maximized germplasm improvement in applied breeding programs. To stay current with applied breeding objectives, it is imperative that necessary adjustments be made as indicated with time. Adjustments should not be made frivolously to cause disruptions in genetic gain from selection. Adjustments made in a haphazard manner would tend to put the improvement program on a treadmill, causing either inconsistent or no genetic improvement for the germplasm chosen to provide materials for the applied breeding program.
Except for mass selection, each recurrent selection scheme includes three operational stages for each cycle of selection:

1. Development of progenies for evaluation.
2. Evaluation of progenies in replicated trials.
3. Recombination of superior progenies chosen on the basis of replicated trials.

Figure 6.1 clearly indicates that this could be directly included in programs developing inbred lines and hybrids by incorporating top progenies of improved cycles into inbred line development programs. For instance, the NDSU maize breeding program has two types of replicated trials (a distinction not commonly clear):

(a) Intra- and inter-population recurrent selection trials for early-maturing germplasm improvement.
(b) Hybrid testcross and single-cross trials for combining ability of inbred lines (see Chapter 1)

Figure 12.13 shows an example on how a breeding program conducts trials classified based on maturity and environmental conditions (e.g., drought to the west, cold and very early maturity to the north). NDSU is a genetic provider as any other seed company. Therefore, lines are coded into foundation seed companies that provide access of them to seed retailers selling hybrids to North Dakota farmers. Lines are also released as parental lines and/or for breeding purposes.

Contrary to choices of populations for improvement and of recurrent selection schemes, the decisions are not easy. Choices of populations and recurrent selection schemes are determined to a large extent on how they relate to the applied breeding program and experience. Very little experience and theory are available to assist in
making decisions regarding operational steps involved in recurrent selection. For instance, how many progenies should be tested to adequately represent genetic variability of the population under selection? How many progenies should be selected for recombination? How many generations of recombination should be conducted before initiating the next cycle of selection? And what method of recombination should be used when 20–30 selections are chosen for recombination? Compromises usually are necessary.

Obviously, the greater the number of populations and progenies available for test the greater the opportunities for having a more representative sample of each population and more chances of success (e.g., alternative heterotic patterns, release of inbred lines, identification of high-yielding hybrids with lowest grain moisture at harvest and fastest dry down). It depends on resources available and how to best allocate them. We need to narrow the choice of populations based on known sources. But testing of progenies in replicated yield trials prevents including a large number intra- or inter-population programs. Most recurrent selection schemes have included a minimum of 100 progenies, but some programs have included 500 or more progenies. Evidence from genetic variance estimates (see Chapter 5) shows that reasonable standard errors are obtained with 100 progenies. However, the greater the progeny number the more valid the estimates. Marquez-Sanchez and Hallauer (1970a, b) found from empirical studies that 180–200 individuals should be included for stabilizing standard errors on estimates of components of variance from use of the nested (design I) mating design. Practically, the breeder should look at the total number of rows possible based on space, labor, and resources and make a decision. If 10,000 rows are possible should the breeder grow 50 populations of 200 individuals or 200 populations of 50 individuals each?

Sample size to be used for recurrent selection schemes is determined to some extent by selection intensity or the number of progenies selected for recombination. If 10% selection intensity is used and 100 progenies are tested, only 10 progenies are recombined to form the next cycle population. Use of only 10 progenies for recombination, however, results in an appreciable level of inbreeding after several cycles of selection (Chapter 9). Empirical evidence agrees with projected effects of inbreeding because of limited number of progenies recombined (Smith, 1979). Some theoretical studies also provide some guidelines for sample sizes to use in plant selection programs. Robertson (1960) showed that expected total advance and half-life of recurrent selection programs were proportional to effective population size. Hence effective population size should be as large as possible for long-term selection programs. Baker and Curnow (1969) and Rawlings (1970) concluded that with a selection intensity of 10%, effective population sizes of 30–45 would be reasonable compromises for both short- and long-term objectives of plant selection programs. The number of progenies to include for recombination depends on the type of progeny for the minimum effective size based on theory. Recent reports have shown the advantages of lowering the effective population size to even 10 progenies though. Weyhrich et al. (1998b) intermated 5, 10, 20, and 30 selected S1 progenies with a common selection intensity of 20% for four to five cycles of S1 progeny selection. Grain yield increased significantly for programs that intermated 10, 20,
and 30 individuals, but grain yield decreased significantly intermitting five individuals. Their data suggested no fewer than 10 should be intermated. Therefore, it seems 15–20 progenies are a minimum, but greater numbers are desirable. Changing number of progenies for recombination changes selection intensity. If 100 progenies are tested, 10% selection intensity (proportion of individuals selected) includes 10 for recombination. If it is desired to increase the number of progenies to say 20 for recombination for 100 progenies tested, selection intensity is reduced to 20%. Then, the selection intensity is reduced when the proportion of individuals selected increases. To maintain selection intensity of 10% by recombining 20 progenies, number of progenies tested needs to be increased to 200. It is desirable to have an adequate sample of progenies for testing and also to have an acceptable selection intensity to realize progress from selection. Obviously some compromises must be made to conduct a manageable selection program. Increasing the selection intensity is desirable for short-term programs and for certain traits but would not be desirable for long-term programs because the effects of inbreeding and genetic drift would hinder future progress. For long-term programs, it seems that at least 20–30 progenies should be included for recombination. Hence, selection intensity is reduced if a smaller number of progenies are tested or increased if more progenies are tested. If, in addition, we use inbred progenies for recombination types of progenies recombined also have an effect, e.g., recombining S1 vs. S2 progenies (see Chapter 9). The inclusion of 20–30 selections for recombination poses problems of the most appropriate method of recombination. If 10 selections are recombined, a diallel mating design frequently has been used to form the next cycle population. Making diallel crosses among 10 selections is not demanding (45 possible crosses and about 90 nursery rows), but as the number of selections increases to 20–30, the number of crosses increases rapidly (see Chapter 4). For example, diallel crosses of 20 selections result in 190 crosses (or possibly 380 nursery rows to make the crosses) as well as the mechanics of making all the crosses. Diallel crossing is feasible if number of selections is 10 or less but becomes increasingly difficult for a greater number of selections. The primary objective of recombination in cyclical selection programs is to intermate superior selections to form the next cycle population for continued selection. Hence, an effective method of recombination is imperative.

Choice of recombination method depends on number of selections included, seasons available, and resources available. The breeder has options:

1. **Diallel mating scheme**: As stated above it is an acceptable method if the number of selections included for recombination is not too large. The advantage is that diallel crossing ensures mating each selection with all other selections.

2. **Chain crossing**: Crossing the selections requires less space than the diallel mating scheme, but equal mating may not be made among selected progenies. Chain crossing 20 selections would require 20 rows in the nursery.

3. **Bulk entry method**: It entails planting selections in alternate rows and remaining rows are planted to a bulk of equal number of seeds of all progenies except the one they are to be crossed with. That is, if row 1 involves selection A, row 2 includes a bulk of all selections except selection A. All plants in row
are crossed with row 1. This method of recombination permits intercrossing among all entries unless some of the plants in the rows that include the bulks are missing. Reciprocal crosses between pairs of rows will minimize the possibility of not making crosses among all selections. If 15 selections are intercrossed, 30 rows are needed to complete the recombination (Table 12.16). In addition to requiring fewer rows than the diallel mating scheme, a great advantage is selfing selections at the same time of recombining. Therefore, for instance, the NDSU maize breeding program utilized bulk-rows as females and selections across recurrent selection programs as males.

Table 12.16 Example of bulk-entry recombination of an intra-population recurrent selection program at NDSU

<table>
<thead>
<tr>
<th>Row No.</th>
<th>Pedigreeb</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>095221</td>
<td>NDSAB(MER-FS)C15 Syn1-106b</td>
<td>074061-106</td>
</tr>
<tr>
<td>095222</td>
<td>BULK 1</td>
<td></td>
</tr>
<tr>
<td>095223</td>
<td>NDSAB(MER-FS)C15 Syn1-5</td>
<td>074061-5</td>
</tr>
<tr>
<td>095224</td>
<td>BULK 2</td>
<td></td>
</tr>
<tr>
<td>095225</td>
<td>NDSAB(MER-FS)C15 Syn1-35</td>
<td>074061-35</td>
</tr>
<tr>
<td>095226</td>
<td>BULK 3</td>
<td></td>
</tr>
<tr>
<td>095227</td>
<td>NDSAB(MER-FS)C15 Syn1-46</td>
<td>074061-46</td>
</tr>
<tr>
<td>095228</td>
<td>BULK 4</td>
<td></td>
</tr>
<tr>
<td>095229</td>
<td>NDSAB(MER-FS)C15 Syn1-83</td>
<td>074061-83</td>
</tr>
<tr>
<td>095230</td>
<td>BULK 5</td>
<td></td>
</tr>
<tr>
<td>095231</td>
<td>NDSAB(MER-FS)C15 Syn1-53</td>
<td>074061-53</td>
</tr>
<tr>
<td>095232</td>
<td>BULK 6</td>
<td></td>
</tr>
<tr>
<td>095233</td>
<td>NDSAB(MER-FS)C15 Syn1-4</td>
<td>074061-4</td>
</tr>
<tr>
<td>095234</td>
<td>BULK 7</td>
<td></td>
</tr>
<tr>
<td>095235</td>
<td>NDSAB(MER-FS)C15 Syn1-99</td>
<td>074061-99</td>
</tr>
<tr>
<td>095236</td>
<td>BULK 8</td>
<td></td>
</tr>
<tr>
<td>095237</td>
<td>NDSAB(MER-FS)C15 Syn1-33</td>
<td>074061-33</td>
</tr>
<tr>
<td>095238</td>
<td>BULK 9</td>
<td></td>
</tr>
<tr>
<td>095239</td>
<td>NDSAB(MER-FS)C15 Syn1-107</td>
<td>074061-107</td>
</tr>
<tr>
<td>095240</td>
<td>BULK 10</td>
<td></td>
</tr>
<tr>
<td>095241</td>
<td>NDSAB(MER-FS)C15 Syn1-30</td>
<td>074061-30</td>
</tr>
<tr>
<td>095242</td>
<td>BULK 11</td>
<td></td>
</tr>
<tr>
<td>095243</td>
<td>NDSAB(MER-FS)C15 Syn1-96</td>
<td>074061-96</td>
</tr>
<tr>
<td>095244</td>
<td>BULK 12</td>
<td></td>
</tr>
<tr>
<td>095245</td>
<td>NDSAB(MER-FS)C15 Syn1-80</td>
<td>074061-80</td>
</tr>
<tr>
<td>095246</td>
<td>BULK 13</td>
<td></td>
</tr>
<tr>
<td>095247</td>
<td>NDSAB(MER-FS)C15 Syn1-102</td>
<td>074061-102</td>
</tr>
<tr>
<td>095248</td>
<td>BULK 14</td>
<td></td>
</tr>
<tr>
<td>095249</td>
<td>NDSAB(MER-FS)C15 Syn1-81</td>
<td>074061-81</td>
</tr>
<tr>
<td>095250</td>
<td>BULK 15</td>
<td></td>
</tr>
</tbody>
</table>

a Row number includes ‘09’ = year 2009 and the actual nursery row number
b Pedigree is based on 12 cycles of modified ear-to-row (MER) selection plus three cycles of full-sib selection, recombining selections to form C16. Syn1 corresponds to one recombination. Currently, there is a program with one and two recombination events on this population as this is Syn1
(4) Equal quantities of seeds of the selections could be bulked and intermatings made among plants of the bulk plantings either by open pollination in isolation plantings or by hand pollination in the breeding nursery. Space would be minimized, but if one or more selections had late flowering, poor germination, or poor competitive ability, some of the selections may not be included equally in the intercrosses.

All four methods of recombination have some disadvantages and advantages, and it is necessary to make a choice that one feels ensures adequate intercrossing among selected progenies. Choice of method used for recombination is not as critical as some other steps in recurrent selection. Different methods of recombination may be used between different cycles of selection without seriously jeopardizing progress of selection.

After initial intercrosses, it is necessary to include a representative sample of plants for additional intercrossing or population maintenance. It is important that duplicate samples be obtained to minimize chances of losing the intercrossed seed because of a crop failure. In some recurrent selection schemes, only one intercrossing is made between cycles of selection. It is imperative that a representative sample be obtained from intercrossed selected progenies to ensure that each one is equally included in bulk samples. The best procedure is to have an equal quantity of seed for each cross by counting the number of ears available and taking an equal number of seeds from each ear to form at least two balanced bulks (Table 12.17). For example, if 10 ears are available for each of 20 crosses and two 1,000 seed samples are needed, five seeds are sampled from each ear for each of the two bulks. If each selection crossed with a bulk planting of the other selections is used for recombination, one needs to count the total number of ears to determine how many seeds per ear should be included in each of the two bulk samples.

For recurrent selection schemes that permit two cycles of recombination between cycles of selection, the same procedures are used to form the bulk samples for the next generation of recombination. One bulk sample is planted either in an isolation field or the nursery for intercrossing and the other bulk sample is placed in suitable cold storage conditions for either emergency or future use. Adequate isolation (at least 200 m and/or significant maturity differences among fields) is necessary to minimize the chances of contamination from other maize plants, and it is easier to obtain in the field than in a nursery planting. If a nursery planting is used for recombination, it is desirable that each plant be included as either a male or female. This requires checking the planting each day to cover silks and prepare tassels and ear shoots for pollination. The usual procedure is to pollinate each ear shoot and not use pollen from each tassel for more than two pollinations or even detasseling and uncovering shoots after pollinating both plants used as female and male so that they are used only once. Pollen also can be bulked and used to pollinate all ear shoots that were prepared for pollination the previous day. Regardless of the system used for the second recombination, equal quantities of seed should be taken from each ear to form at least two bulk samples of the Synthetic 2 seed, one for cold storage and the second to initiate the next cycle of selection. At least 500 seeds per bulk would
Table 12.17 Inventory for the 2008 population maintenance section of the NDSU maize breeding program generates three balanced bulks per population

<table>
<thead>
<tr>
<th>Inventory No.</th>
<th>Pedigree</th>
<th>2007 source</th>
<th>Weight (g)</th>
<th>No. of Ears</th>
</tr>
</thead>
<tbody>
<tr>
<td>087001-7020-1B</td>
<td>NDSS Syn 1</td>
<td>077000-7039-2</td>
<td>185</td>
<td>143</td>
</tr>
<tr>
<td>087001-7020-2B</td>
<td>NDSS Syn 1</td>
<td>077000-7039-2</td>
<td>185</td>
<td>143</td>
</tr>
<tr>
<td>087001-7020-3B</td>
<td>NDSS Syn 1</td>
<td>077000-7039-2</td>
<td>185</td>
<td>143</td>
</tr>
<tr>
<td>087021-7040-1B</td>
<td>NDCG(FS)C1 Syn 3</td>
<td>077161-7170-2</td>
<td>168</td>
<td>177</td>
</tr>
<tr>
<td>087021-7040-2B</td>
<td>NDCG(FS)C1 Syn 3</td>
<td>077161-7170-2</td>
<td>167</td>
<td>177</td>
</tr>
<tr>
<td>087021-7040-3B</td>
<td>NDCG(FS)C1 Syn 3</td>
<td>077161-7170-2</td>
<td>169</td>
<td>177</td>
</tr>
<tr>
<td>087041-7060-1B</td>
<td>NDL Syn 1</td>
<td>077041-7080-2</td>
<td>172</td>
<td>125</td>
</tr>
<tr>
<td>087041-7060-2B</td>
<td>NDL Syn 1</td>
<td>077041-7080-2</td>
<td>171</td>
<td>125</td>
</tr>
<tr>
<td>087041-7060-3B</td>
<td>NDL Syn 1</td>
<td>077041-7080-2</td>
<td>172</td>
<td>125</td>
</tr>
<tr>
<td>087061-7080-1B</td>
<td>NDBSK(H-M)C3 Syn 1</td>
<td>077081-100-1</td>
<td>208</td>
<td>136</td>
</tr>
<tr>
<td>087061-7080-2B</td>
<td>NDBSK(H-M)C3 Syn 1</td>
<td>077081-100-1</td>
<td>206</td>
<td>136</td>
</tr>
<tr>
<td>087061-7080-3B</td>
<td>NDBSK(H-M)C3 Syn 1</td>
<td>077081-100-1</td>
<td>205</td>
<td>136</td>
</tr>
<tr>
<td>087081-7100-1B</td>
<td>NDBS11(FR-M)C3 Syn 1</td>
<td>077101-20-1</td>
<td>195</td>
<td>122</td>
</tr>
<tr>
<td>087081-7100-2B</td>
<td>NDBS11(FR-M)C3 Syn 1</td>
<td>077101-20-1</td>
<td>190</td>
<td>122</td>
</tr>
<tr>
<td>087081-7100-3B</td>
<td>NDBS11(FR-M)C3 Syn 1</td>
<td>077101-20-1</td>
<td>192</td>
<td>122</td>
</tr>
<tr>
<td>087101-7120-1B</td>
<td>NDBS1011 Syn 1</td>
<td>077121-40-1</td>
<td>181</td>
<td>127</td>
</tr>
<tr>
<td>087101-7120-2B</td>
<td>NDBS1011 Syn 1</td>
<td>077121-40-1</td>
<td>178</td>
<td>127</td>
</tr>
<tr>
<td>087101-7120-3B</td>
<td>NDBS1011 Syn 1</td>
<td>077121-40-1</td>
<td>179</td>
<td>127</td>
</tr>
<tr>
<td>087121-7140-1B</td>
<td>NDBS1011(FR-M-FS)C5 Syn 1</td>
<td>075699-5722-1</td>
<td>181</td>
<td>141</td>
</tr>
<tr>
<td>087121-7140-2B</td>
<td>NDBS1011(FR-M-FS)C5 Syn 1</td>
<td>075699-5722-1</td>
<td>176</td>
<td>141</td>
</tr>
<tr>
<td>087121-7140-3B</td>
<td>NDBS1011(FR-M-FS)C5 Syn 1</td>
<td>075699-5722-1</td>
<td>179</td>
<td>141</td>
</tr>
<tr>
<td>087141-7160-1B</td>
<td>BS22(R-T1)C9 Syn 1</td>
<td>075501-24-1</td>
<td>183</td>
<td>136</td>
</tr>
<tr>
<td>087141-7160-2B</td>
<td>BS22(R-T1)C9 Syn 1</td>
<td>075501-24-1</td>
<td>182</td>
<td>136</td>
</tr>
<tr>
<td>087141-7160-3B</td>
<td>BS22(R-T1)C9 Syn 1</td>
<td>075501-24-1</td>
<td>183</td>
<td>136</td>
</tr>
<tr>
<td>087161-7180-1B</td>
<td>[NDSAB(MER-FS)C15 Syn 2]Syn 1</td>
<td>075574-96-1</td>
<td>218</td>
<td>122</td>
</tr>
<tr>
<td>087161-7180-3B</td>
<td>[NDSAB(MER-FS)C15 Syn 2]Syn 1</td>
<td>075574-96-1</td>
<td>215</td>
<td>122</td>
</tr>
</tbody>
</table>
Table 12.17 (continued)

<table>
<thead>
<tr>
<th>Inventory No.</th>
<th>Pedigree</th>
<th>2007 source</th>
<th>Weight (g)</th>
<th>No. of Ears</th>
</tr>
</thead>
<tbody>
<tr>
<td>087181-7200 -1B</td>
<td>BS22(R-T2)C9 Syn1</td>
<td>075526~72-1</td>
<td>184</td>
<td>139</td>
</tr>
<tr>
<td>087181-7200 -2B</td>
<td>BS22(R-T2)C9 Syn1</td>
<td>075526~72-1</td>
<td>182</td>
<td>139</td>
</tr>
<tr>
<td>087181-7200 -3B</td>
<td>BS22(R-T2)C9 Syn1</td>
<td>075526~72-1</td>
<td>183</td>
<td>139</td>
</tr>
<tr>
<td>087201-7220 -1B</td>
<td>[NDSM(M-FS)C9 Syn2]Syn1</td>
<td>075622~48-1</td>
<td>194</td>
<td>119</td>
</tr>
<tr>
<td>087201-7220 -2B</td>
<td>[NDSM(M-FS)C9 Syn2]Syn1</td>
<td>075622~48-1</td>
<td>191</td>
<td>119</td>
</tr>
<tr>
<td>087201-7220 -3B</td>
<td>[NDSM(M-FS)C9 Syn2]Syn1</td>
<td>075622~48-1</td>
<td>190</td>
<td>119</td>
</tr>
<tr>
<td>087221-7240 -1B</td>
<td>BS21(R-T)C9 Syn1</td>
<td>075598~620-1</td>
<td>173</td>
<td>137</td>
</tr>
<tr>
<td>087221-7240 -2B</td>
<td>BS21(R-T)C9 Syn1</td>
<td>075598~620-1</td>
<td>173</td>
<td>137</td>
</tr>
<tr>
<td>087221-7240 -3B</td>
<td>BS21(R-T)C9 Syn1</td>
<td>075598~620-1</td>
<td>171</td>
<td>137</td>
</tr>
<tr>
<td>087241-7260 -1B</td>
<td>CGL(S-FR2)C1 Syn1</td>
<td>075674~97-1</td>
<td>177</td>
<td>155</td>
</tr>
<tr>
<td>087241-7260 -2B</td>
<td>CGL(S-FR2)C1 Syn1</td>
<td>075674~97-1</td>
<td>174</td>
<td>155</td>
</tr>
<tr>
<td>087241-7260 -3B</td>
<td>CGL(S-FR2)C1 Syn1</td>
<td>075674~97-1</td>
<td>173</td>
<td>155</td>
</tr>
<tr>
<td>087261-7280 -1B</td>
<td>BS21(R-FR2)C1 Syn1</td>
<td>075650~72-1</td>
<td>160</td>
<td>130</td>
</tr>
<tr>
<td>087261-7280 -2B</td>
<td>BS21(R-FR2)C1 Syn1</td>
<td>075650~72-1</td>
<td>159</td>
<td>130</td>
</tr>
<tr>
<td>087261-7280 -3B</td>
<td>BS21(R-FR2)C1 Syn1</td>
<td>075650~72-1</td>
<td>160</td>
<td>130</td>
</tr>
<tr>
<td>087281-7300 -1B</td>
<td>CGL(S-FR3)C1 Syn1</td>
<td>075751~70</td>
<td>149</td>
<td>132</td>
</tr>
<tr>
<td>087281-7300 -2B</td>
<td>CGL(S-FR3)C1 Syn1</td>
<td>075651~70</td>
<td>147</td>
<td>132</td>
</tr>
<tr>
<td>087281-7300 -3B</td>
<td>CGL(S-FR3)C1 Syn1</td>
<td>075651~70</td>
<td>147</td>
<td>132</td>
</tr>
<tr>
<td>087301-7320 -1B</td>
<td>BS21(R-FR3)C1 Syn1</td>
<td>075772~90-1</td>
<td>167</td>
<td>135</td>
</tr>
<tr>
<td>087301-7320 -2B</td>
<td>BS21(R-FR3)C1 Syn1</td>
<td>075772~90-1</td>
<td>165</td>
<td>135</td>
</tr>
<tr>
<td>087301-7320 -3B</td>
<td>BS21(R-FR3)C1 Syn1</td>
<td>075772~90-1</td>
<td>165</td>
<td>135</td>
</tr>
<tr>
<td>087321-7340 -1B</td>
<td>NDSHLC(FS-M-FS)C5 Syn1</td>
<td>075725~49-1</td>
<td>182</td>
<td>97</td>
</tr>
<tr>
<td>087321-7340 -2B</td>
<td>NDSHLC(FS-M-FS)C5 Syn1</td>
<td>075725~49-1</td>
<td>178</td>
<td>97</td>
</tr>
<tr>
<td>087321-7340 -3B</td>
<td>NDSHLC(FS-M-FS)C5 Syn1</td>
<td>075725~49-1</td>
<td>184</td>
<td>97</td>
</tr>
<tr>
<td>087341-7360 -1B</td>
<td>EARLYGEM</td>
<td>075791~5823-1</td>
<td>186</td>
<td>123</td>
</tr>
<tr>
<td>087341-7360 -2B</td>
<td>EARLYGEM</td>
<td>075791~5823-1</td>
<td>182</td>
<td>123</td>
</tr>
<tr>
<td>087341-7360 -3B</td>
<td>EARLYGEM</td>
<td>075791~5823-1</td>
<td>183</td>
<td>123</td>
</tr>
</tbody>
</table>
be advisable. When recombination is done through controlled hand pollination, there is an opportunity to control the number of female and male gametes that are contributed by each female and male parent to the next generation. Therefore, taking a sample of equal size from each hand-pollinated ear we will get the highest effective number for a given number of progenies in the recombination block (see Chapter 11).

A simpler procedure that can be used to recombine selected progenies does not involve controlled hand pollination. Seeds from each selected progeny are planted ear-to-row, and rows are detasseled before flowering. A bulk of seeds of all progenies is planted in alternate rows as males, which will provide pollen for all detasseled rows so that each female row is pollinated with a bulk of pollen of all progenies. The ratio male:female rows may vary from 1:1 to 1:4 and male rows can be split planted to reduce assortative mating. Because pollination is random there is no control on the number of male gametes contributed by each male plant, but a control is possible on the number of female gametes by taking a sample of equal size from each pollinated ear. This sampling procedure assures a larger effective population size than would result if a single sample were taken from a bulk of the whole set of seeds. On the other hand, control of the number of female gametes will only result in a smaller effective population size as compared with other recombination procedures previously discussed. Small differences in effective population size make no great difference when number of progenies saved for recombination is not too small and/or inbred.

Effects of linkage are a concern when considering recombination number of between cycles of selection. Effects of linkage in recurrent selection programs are not as obvious as effects of inbreeding. Linkage effects have been examined in F2 populations of two inbred lines and they did affect estimates of genetic components of variance (see Chapter 5). Progenies were formed from individuals in F2 populations for estimating genetic parameters. Broad-based populations that have had several opportunities for recombination would be expected to be near linkage equilibrium. Difficulty of linkage effects arises from recombination of selected progenies. Because linkage is known to exist, the question is how many generations of recombination are needed for the selected population to approach linkage equilibrium. Hanson (1959) concluded that four recombination generations seemed to be adequate. In temperate zones four generations of intermating would require 2 years to complete, provided winter nurseries were available. A minimum number (one or two) of intermating generations are used in most recurrent selection programs. In some recurrent selection schemes (e.g., Table 12.7), an additional generation of recombination can be included without increasing cycle time if efficient use is made of winter nurseries. Additional recombination of selected progenies is usually acknowledged to be useful. Without any empirical evidence at the moment, it seems that greater gains can be made by increasing the number of cycles of selection than by increasing the number of recombination generations between cycles of selection.

Latter (1965, 1966a, b) examined effects of linkage on limits of selection. He concluded that reduction in total response was minor unless recombination probability among loci was less than 0.1. As interpreted by Rawlings (1970), if recombination probability was 0.1 reduction was about 2.5% and if recombination
probability was 0.05 reduction was about 6%. Linkage had very little effect on rate of response in early generations of selection, particularly if the populations were in linkage equilibrium when selection was started. Rawlings concluded that if linkage effects are a concern, effective population size should be increased to allow for linkage depression. It seems that linkage effects will not seriously affect response to selection although the breeding might want to approach linkage equilibrium when improving populations. A different approach could be followed when developing inbred lines. In most breeding programs developing of inbred lines is within heterotic groups in order to have a greater expression of heterosis in the final hybrids. However, breeders can, in certain occasions, develop inbred lines from heterotic patterns directly. The NDSU maize breeding program develops, in 5% of the cases (see Chapter 1), lines from hybrids that have demonstrated specific genetic effects to be outstanding in which presumably linkage blocks are well defined and it is desirable to keep them. The public example would be B73 × Mo17.

The basic relation between factors involved in operational stages of recurrent selection was given by Eberhart (1970) and Sprague and Eberhart (1977). They expressed predicted gain $\Delta G$ in the following relation:

$$\Delta G = \frac{ck\hat{\sigma}^2_g}{y\left[\hat{\sigma}^2/(re) + \hat{\sigma}^2_{ge}/e + \hat{\sigma}^2_g\right]^{1/2}}$$

where $c$ is the parental control (see Chapter 6), $k$ the standardized selection differential, $\hat{\sigma}^2_g$ the portion of genetic variance among progenies due to additive effects, $y$ the number of years per cycle, and $\left[\hat{\sigma}^2/(re) + \hat{\sigma}^2_{ge}/e + \hat{\sigma}^2_g\right]$ the phenotypic variance. This formula differs slightly from those in Chapter 6, where $k$ in the numerator is selection intensity (which shows if selection is for one or both sexes or parental control) and $c$ is a coefficient that takes into account the crossing system (which dictates the covariance between relatives in different generations); $k$ and $c$ together multiply $\hat{\sigma}^2_A$ (additive genetic variance) or one of its homologues. On the other hand, in the formula presented by Sprague and Eberhart (1977), $k$ has the same meaning and $c$ is a coefficient that takes into account only parental control. The coefficient that is a result from the covariance between relatives is inherently contained in $\hat{\sigma}^2_g$, which is the additive portion of variance expressed among selection units or the portion of covariance between relatives in the same generation. For example, the numerator for expected gain from selection among half-sib families for only one sex (ear-to-row procedure) is shown in Chapter 6 as being $k(\frac{1}{8})\hat{\sigma}^2_A$, where $\frac{1}{8}$ is the coefficient for the covariance between half-uncle–nephew. On the other hand, Sprague and Eberhart (1977) give the same coefficient expressed as $k(\frac{1}{2})\hat{\sigma}^2_A$ where $\frac{1}{2}$ is a coefficient due to parental control (selection for one sex) and $\hat{\sigma}^2_A$ is the additive portion of variance among half-sib families or covariance between half-sibs (same generation). Although both formulations give essentially the same results, a discussion of differences between them follows. Genetic gain in the formula is expressed on a per-year basis. Relation for predicted gain can assist in making decisions in each operational stage of recurrent selection. Table 7.35 shows that the prediction formula provides
guidelines in the choice of recurrent selection procedure that maximizes rate of gain per year. We want to obtain maximum gain possible per year. If the choice among recurrent selection schemes is on a genetic gain per-year basis, the choice may be different if different numbers of years are required to complete each cycle of selection (Table 7.35). Prediction formulas for different methods of recurrent selection are listed in Table 12.18.

The predicted gain formula shows that genetic gain can be increased if parental control is increased, years per cycle are reduced, and additive genetic variance among progenies is increased. Two other important factors are selection intensity and phenotypic variance. If information is available, various combinations of these five factors can be determined to provide some guidelines for the potentially most efficient recurrent selection scheme to use (Empig et al., 1972). The formula does not provide any guidelines on choice of population. The three operational stages of recurrent selection can be adjusted to affect predicted gain. The number of progenies to test and recombine affects selection intensity and the number of generations of recombination determines how many years are required to complete each cycle of selection. How the various factors affect efficiency of selection is illustrated in Chapter 7.

Although the prediction formula can provide assistance in determining the most efficient method for conducting recurrent selection, the difference in predicted gain per cycle or year may often be too small to be detectable in experimental trials. Then the appropriate course is to use the recurrent selection scheme that satisfies objectives of the applied breeding program. Other factors that enhance efficiency of recurrent selection schemes include such items as use of mechanical planting and harvesting equipment to permit timely handling of a greater number of plots in the most accurate way (less error), high-speed computers to analyze large volumes of data, adequate test sites to estimate the genotype by environment interaction, and availability of off-season nurseries. These items are not included directly in the formula, but they contribute greatly to predicted gain on a per-year basis. Inclusion of these items permits better sampling to get a better estimate of $\hat{\sigma}_g^2$, reduces number of years per cycle, and reduces confounding effects of genotype by environment interactions. Perhaps, the most important factor that increases the difference between observed and predicted values is genotype by environment interactions. Therefore, identifying methods and mechanisms that can allow more extensive testing per cycle are desirable. It is also encouraged to keep searching for new methodologies that can increase genetic gain of existing recurrent selection methods and/or new ones to come.

It is necessary to conduct cyclical selection programs on a regular basis without interruptions. If the population improvement program via some form of recurrent selection does not contribute to breeding program objectives, it should not be conducted. Basic selection studies are conducted to provide information for applied breeders, but they also provide breeding material for applied programs (e.g., Hallauer, 1973a, b, 1992; Suwantaradon and Eberhart, 1974; Russell and Eberhart, 1975; Hallauer and Carena, 2009).
Table 12.18  Expected genetic gain per cycle ($\Delta \bar{G}$) for different intra- and inter-population methods of recurrent selection

<table>
<thead>
<tr>
<th>Selection method</th>
<th>Expected genetic gain ($\Delta \bar{G}$)</th>
<th>Years per cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mass$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. One sex</td>
<td>$k(\hat{\gamma}<em>2)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}<em>w^2 + \hat{\sigma}</em>{DE}^2 + \hat{\sigma}</em>{AE}^2 + \hat{\sigma}</em>{D}^2 + \hat{\sigma}_A^2}$</td>
<td>1</td>
</tr>
<tr>
<td>b. Both sexes</td>
<td>$k\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}<em>w^2 + \hat{\sigma}</em>{DE}^2 + \hat{\sigma}</em>{AE}^2 + \hat{\sigma}_{D}^2 + \hat{\sigma}_A^2}$</td>
<td>1 or 2</td>
</tr>
<tr>
<td>2. Modified ear-to-row$^{b,c}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. One sex</td>
<td>$k(\hat{\gamma}_8)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \hat{\gamma}<em>3(\hat{\sigma}</em>{AE}^2)} + (\hat{\gamma}_4)\hat{\sigma}_A^2$</td>
<td>1</td>
</tr>
<tr>
<td>b. Both sexes</td>
<td>$k(\hat{\gamma}_8)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \hat{\gamma}<em>3(\hat{\sigma}</em>{AE}^2)} + (\hat{\gamma}_4)\hat{\sigma}_A^2$</td>
<td>2</td>
</tr>
<tr>
<td>3. Half-sib$^d$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Remnant half-sib seed</td>
<td>$k(\hat{\gamma}_4)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \hat{\gamma}<em>3(\hat{\sigma}</em>{AE}^2)} + (\hat{\gamma}_4)\hat{\sigma}_A^2$</td>
<td>2</td>
</tr>
<tr>
<td>b. Self-seed</td>
<td>$k(\hat{\gamma}_2)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \hat{\gamma}<em>3(\hat{\sigma}</em>{AE}^2)} + (\hat{\gamma}_4)\hat{\sigma}_A^2$</td>
<td>3</td>
</tr>
<tr>
<td>4. Full-sib</td>
<td>$k(\hat{\gamma}_2)\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \left[\frac{\hat{\gamma}<em>2\hat{\sigma}</em>{AE}^2 + (\hat{\gamma}<em>4)\hat{\sigma}</em>{DE}^2}{e}\right] + \left[\hat{\gamma}_2\hat{\sigma}_A^2 + (\hat{\gamma}<em>4)\hat{\sigma}</em>{D}^2\right]}$</td>
<td>2</td>
</tr>
<tr>
<td>5. S$^1_e$</td>
<td>$k\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \left[\hat{\gamma}<em>3\hat{\sigma}</em>{AE}^2 + (\hat{\gamma}<em>4)\hat{\sigma}</em>{DE}^2\right] + \left[\hat{\sigma}_A^2 + (\hat{\gamma}_4)\hat{\sigma}_D^2\right]}$</td>
<td>2</td>
</tr>
<tr>
<td>6. S$^2_e$</td>
<td>$k\hat{\gamma}_2\hat{\sigma}<em>A^2/\sqrt{\hat{\sigma}</em>{re}^2 + \left[\hat{\gamma}<em>2\hat{\sigma}</em>{AE}^2 + (\hat{\gamma}<em>4)\hat{\sigma}</em>{DE}^2\right] + \left[\hat{\gamma}_2\hat{\sigma}_A^2 + (\hat{\gamma}_4)\hat{\sigma}_D^2\right]}$</td>
<td>3</td>
</tr>
<tr>
<td>7. Reciprocal recurrent$^f$</td>
<td>$k_1(\hat{\gamma}<em>4)\hat{\sigma}</em>{A_{12}}^2/\sqrt{\hat{\sigma}<em>{re}^2 + \left[\hat{\gamma}<em>3\hat{\sigma}</em>{AE</em>{12}}^2 + (\hat{\gamma}<em>4)\hat{\sigma}<em>A</em>{12}^2\right]} + k_2(\hat{\gamma}<em>4)\hat{\sigma}</em>{A</em>{21}}^2/\sqrt{\hat{\sigma}<em>{re}^2 + \left[\hat{\gamma}<em>2\hat{\sigma}</em>{AE</em>{21}}^2 + (\hat{\gamma}_4)\hat{\sigma}<em>A</em>{21}^2\right]}$</td>
<td>3</td>
</tr>
</tbody>
</table>
### Table 12.18 (continued)

<table>
<thead>
<tr>
<th>Selection method</th>
<th>Expected genetic gain ($\Delta G^a$)</th>
<th>Years per cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Reciprocal full-sib</td>
<td>$\frac{k(\delta_2^2)}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_2^2}{\sigma^2})}{e} \left[ \frac{\left(\frac{\delta_2^2}{\sigma^2}\right)}{RE} + \frac{(\frac{\delta_2^2}{\sigma^2})}{AE} + \frac{(\frac{\delta_2^2}{\sigma^2})}{DE} \right]} + \left(\frac{\delta_2^2}{\sigma^2}\right) \right]}$</td>
<td>2</td>
</tr>
<tr>
<td>9. RRS based on testcross of half-sib families$^f$</td>
<td>$\frac{k_1(\delta_1^2)}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_1^2}{\sigma^2})}{e} + \left(\frac{\delta_1^2}{\sigma^2}\right) \right]} + \frac{k_2(\delta_1^2)}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_1^2}{\sigma^2})}{e} + \left(\frac{\delta_1^2}{\sigma^2}\right) \right]}$</td>
<td>3</td>
</tr>
<tr>
<td>10. RRS based on half-sib families of prolific plants$^f$</td>
<td>$\frac{k_1(\delta_1^2)}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_1^2}{\sigma^2})}{e} + \left(\frac{\delta_1^2}{\sigma^2}\right) \right]} + \frac{k_2(\delta_1^2)}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_1^2}{\sigma^2})}{e} + \left(\frac{\delta_1^2}{\sigma^2}\right) \right]}$</td>
<td>1 or 2</td>
</tr>
</tbody>
</table>

$^a$k is the standardized selection differential; $\delta_A^2$ and $\delta_D^2$ are the additive and dominance variances; $\delta_{AE}^2$ and $\delta_{DE}^2$ are the additive and dominance by environment interactions; $r$ is the number of replications in each of $e$ environments; $\hat{\sigma}_{re}^2$ is within-plot environmental variation; and $\hat{\sigma}^2$ is experimental error.

$^b$Parental control is only for female and one sex and both parents for both sexes. Add $\hat{\sigma}^2$ if no blocking.

$^c$If mass selection is practiced within ear rows for the primary trait, an additional component should be added to the expected gain: $\frac{k(\frac{\delta_2^2}{\sigma^2})}{\sqrt{\frac{\sigma^2}{r} + \frac{(\frac{\delta_2^2}{\sigma^2})}{e} + \left(\frac{\delta_2^2}{\sigma^2}\right) \right]} + \left(\frac{\delta_2^2}{\sigma^2}\right)$

$^d_a$: Use $\delta_A^2$ for parental population tester; $^b$: use $\delta_A^2 = 2pq\alpha_1\alpha_2$ for unrelated tester.

$^e$Definitions of additive genetic variances change slightly with inbreeding. If $p = q = 0.5$ dominance variances are easier to define.

$^f$All $\delta_{A}$ are homologues of $\delta_{A}^2$. $\delta_{A}^2 = (\frac{\delta_{A}^2}{2})(\delta_{A}^2 + \delta_{A}^2)$ is the additive genetic variance in the crossed population.

---

Table continued...
Pedigree and backcross methods of breeding will continue to have a prominent role, particularly in maize breeding programs that utilize and incorporate molecular markers for transfer of single-gene traits. Recurrent selection methods are intended to supplement these breeding methods and should be considered as only one other type of breeding method in the breeder’s arsenal available for developing improved varieties and hybrids. One will not replace the other. The time and effort expended on recurrent selection in applied breeding programs are justified only when they contribute to them. Integration of recurrent selection methods with other selection and breeding methods is feasible in nearly all instances. Often recurrent selection programs are considered within long-term pre-breeding efforts to improve elite germplasm. Therefore, recurrent selection methods have been used significantly in the public sector with the goals to increase the genetic diversity of germplasm available and close the gap between non-elite and useful competitive genetic materials. B73 was a result of the integration of recurrent selection with inbred line development.

Comparisons of recurrent selection and classical pedigree methods of breeding are fortunately limited since these methods are complementary rather than a matter of choice. Sprague (1952) reported comparisons for selection for oil content of the kernel and Duvick (1977) for yield. Depending on the comparison, Sprague reported that recurrent selection was two to five times more effective than pedigree selection and the recurrent selection program still had genetic variability present for continued response to selection. Duvick found that rates of gain were nearly equal for the two distinct methods of breeding; 0.68 q/ha per year of pedigree selection (13.3 years per cycle) vs. 0.71 q/ha per year for recurrent selection (3 years per cycle). The comparisons of Sprague and Duvick do not resolve which method of selection is more effective, and it is doubtful whether any arguments can show that one method (pedigree or recurrent) is more important or more efficient with present techniques and facilities; however, cycle time for pedigree selection has been reduced 50% or more than the 13.3 years reported by Duvick. Pedigree and recurrent selection methods are two different methods of selection with different objectives. Pedigree selection is usually restricted to breeding populations having a narrow genetic base (often developed from a cross of two inbred lines) that emphasizes inbreeding, while recurrent selection methods are applied to populations having a broad genetic base that minimizes effects of inbreeding. Generally, pedigree selection is the most common method for inbred line development which focuses on both qualitative and quantitative traits while recurrent selection improves germplasm sources for inbred line development and works mostly with quantitative traits. The breeding methods are usually quite distinct. Effective recurrent selection methods contribute selections that can be used either in pedigree selection programs or directly as parental stocks in hybrids or as improved varieties; that is, recurrent selection supplements other methods of selection (e.g., Mikel, 2006; Carena and Wicks III, 2006; Hallauer and Carena, 2009).
Jenkins (1978) emphasized the importance of pedigree and backcrossing methods in maize breeding programs for developing inbred lines. In 1936, 350 lines developed by state and federal agencies were listed. Only 7 (2%) of the lines listed were second-cycle lines. Subsequent summaries of inbred lines released by state and federal agencies show that the percentage of second-cycle lines was 20% in 1948, 26% in 1952, 40% in 1956, and 50% in 1960. Since 1960 most of the new inbred lines released by state and federal agencies have been second-cycle lines developed by pedigree and backcrossing methods of selection. The same trend is occurring in proprietary lines. The genetic base of breeding materials, therefore, is becoming more restricted. Recurrent selection methods can contribute to broadening the genetic base of breeding programs, but only when all methods of selection are integrated in all phases of breeding programs will the projected goals of continuous genetic improvement be realized. Early × late crosses and backcrosses (Rinke and Sentz, 1962; Carena et al., 2009b) as well as pedigree selection of diverse improved populations will also be able to significantly improve genetic gains and increase the genetic diversity of hybrids available in the market.

The NDSU maize breeding has 80 years of continued maize breeding research. Its objectives are to adapt exotic elite genetically broad-based germplasm (e.g., NDSU EarlyGEM program, stratified mass selection); maximize genetic improvement of adapted germplasm (e.g., intra- and inter-population recurrent selection programs); develop unique early-maturing lines, populations, and NDSU lines × industry hybrids; and educate plant breeding MS and PhD graduates. This is an example where all breeding techniques discussed can be applied for developing new and unique products. This program is the most northern US public maize breeding program in North America where most industry hybrids are not locally bred (although there is an increased interest with almost three million acres in North Dakota). Therefore, maize industry hybrids are often late-maturing, lower in quality, and do not even offer traits (e.g., drought tolerance) that are needed in areas where the market is not big enough for large investments. The lack of early-maturing industry elite testers is largely part of the problem. Therefore, cooperation is essential for any breeding program targeting the use of all breeding methods discussed. Access to non-obsolete technology and cooperation (SNP markers, doubled haploids, grain quality screening, managed stress environments and other winter nurseries, plot exchanges, etc.) is encouraged to keep plant breeding programs strong, especially when deciding between accessing state of the art industry labs vs. investing with federal grants in academic labs that become obsolete very quickly (e.g., a particular molecular project might take 1 week in industry while 2 years in the academic lab). It has been emphasized how important is the use of winter generations to speed up genetic improvement. Winter nurseries allowing more than two seasons per year have provided unique <90RM NDSU EarlyGEM lines in 4 years in addition to inbred and hybrid screening for drought tolerance and seed for hybrid testing every year (Carena et al., 2009a, b).

While there have been successful innovations in US public and private maize breeding during the past 100 years (e.g., the public inbred–hybrid concept, the scientific basis of plant breeding that provided a genetic basis for variation and
valid experimental techniques for measuring differences, genetic diversity and B73, biotechnology) public US maize breeding programs have either discontinued or are eroding because of decreasing state and federal funding toward basic science (e.g., genome sequence of B73, QTL-mapping strategies). One thing that is certain is that there is no limit to genetic improvement when most genes are targeted in the breeding process. Heterotic effects are unique for each hybrid and sequencing efforts on only B73 and related materials may limit the discovery of useful alleles for complex traits in tropical, subtropical, and early-maturing genetic materials.

References


References


Hanson, W. D. 1959. The breakup of initial linkage blocks under selected mating systems. Genetics 44:857–68.


Noble, S. W. 1966. Sampling of heterogeneous testers and the comparison of a double cross with parental and non-parental single crosses as testers for the evaluation of maize lines. Ph.D. dissertation, Iowa State University, Ames, IA.


Index

A
Additive variance
defined, 47, 50, 52
estimates/estimation, 57–58, 61–65, 170–175, 179
prediction, 268
Allele
fixation, 262
substitution, 44–45
Alph, 186, 320–321, 323–324, 347
Androgenesis, 11, 24–25
Andropogoneae, 535
Avena, 122

B
Beta vulgaris, 383
Biparental progenies, 93
Breeding
hybrid, 10, 64, 226–227, 384, 487, 501
objectives, 532, 634
plans, 9, 47, 260, 577–649
value, 44–46, 49, 57, 70, 228, 231, 279, 281, 384
Burr’s White, 5, 314, 429

C
Cateto, 266, 296, 306, 310, 342–343, 565
Chionachne, 535
Clarage, 315
Coix, 535–535
Combining ability
general, 123, 125, 205, 227, 231, 318, 321, 333–334, 383, 440, 446, 518, 600, 609–610, 612, 620–622, 630
Corn Belt Dents, 2, 6, 546, 548–549, 559, 564, 566, 568
Corn stunt, 329
Correlations
early vs. late testing, 385
line vs. hybrid, 413
phenotypic, 211, 215, 277–279, 306, 400
among traits, 210–216, 282, 402, 587


D
Dawes, 319, 323
Design I
analysis, 96–97, 99, 101
description, 160
estimates, 96–99, 101
Design II
analysis, 105, 107, 109
description, 103
estimates, 104–105
Design III
analysis, 115–116
description, 80, 113
estimates, 113–118
Diabrotica, 631
Diallel
analysis, 105, 123, 125–131, 133, 139, 141, 143, 162, 170, 416, 505
covariances of relatives, 97, 101, 109, 126, 135, 141, 162, 199
description, 87, 93, 105
general combining ability, 622
partial, 97, 133–137, 508, 613
specific combining ability, 622
Diplodia zea, 329, 583–584
Diversity
heterosis, 478, 487–489, 520, 522
selection, 410
Dominance
E
Ear-to-row selection
description, 34, 245
experimental results, 15
theory, 14
Earworm, 311, 322
Ecuador, 573, 343
Effect
average, 43–47, 140, 325, 392–393, 397
gene substitution, 44, 397
Effective population size, 234, 261–269, 271–272, 276, 292, 300
Endosperm, 25, 235–236, 426, 487, 538–539, 543, 545
Epistasis
effects, 179, 454
estimates/estimation, 60, 128, 162, 164, 199, 202, 206, 462
Estimates, experimental
correlations, 200, 210–216
genetic effects, 206–207, 209
heterosis, 297, 488–489, 495, 513, 519–521
inbreeding, 180, 199, 204–205, 207–208, 213
variance, 169–216
ETO Composite, 197, 300–301, 455, 561–562
Euchlaena, 534
European corn borer, 328, 359–364, 433, 455, 583–585, 587, 589, 596–598, 621, 624
F
Flint, 2, 349, 407, 486–487, 491, 496, 538–539, 545–548, 559, 561, 564–566, 602, 606, 613, 630
Flint Composite, 309–310, 317, 327, 342–343, 356, 613
Full-sib
family means, 89, 113, 250
inbreeding, 39–42, 52
reciprocal recurrent selection, 350–351, 454
selfing, 14, 17, 20, 26
variance, 72–73, 132
Gardner-Eberhart model, 130–131
Gene
  average effect, 44–45, 397
  expected change, 263, 269, 271, 395
  fixation, 435–436
  frequency, 394, 549
  substitution, 44–45, 397
General combining ability, see Combining ability
Generation means, 60–63, 65, 205, 451–452, 470
Genetic
  coefficient of variability, 148, 268, 291, 353, 355–357
  uniformity, 534
  variance, see Variance
  vulnerability, 229, 531, 534, 550–552
Genotypic frequency, 34–35, 149
Genotypic value, 43, 36–39, 41–43, 45, 47–48, 50, 70, 149, 152, 237, 284, 386, 499
Germplasm
  classification, 548
  exotic, 8, 20, 197, 300, 304
  preservation, 550, 553–554
Glycine
  122
Haploids, 23–24, 146, 385, 413, 470, 577, 648
Half-sib
  description, 603, 606
  ear-to-row, 383
  means, 11
  selection, see Selection
testcross, 605
  theory, 34, 41
Heliothis zea, 311
Helminthosporium
  maydis, 10
  turcicum, 206, 311
Heritability
Heterosis
  biometrical, 497, 499
  components of means, 515
  dominant favorable factors, 494
  epistasis, 10
  estimates, 116
  genetic basis, 493, 495, 519, 522, 580
  heterozygosity per se, 494
  high parent, 344–345, 480–481, 486, 519, 522
  mid parent, 335, 339–341, 351, 481, 486, 488, 519–521
  overdominance, 119, 373, 388, 394, 417
  patterns, 224, 256, 292, 494
  physiological stimulation, 494, 560–561
  prediction methods, 499–515
  specific, 131, 224, 489, 515, 518
  variety, 131, 224, 489, 515, 517–518
Hickory King, 532
Hogue’s Yellow Dent, 304, 315, 427
Homozygous diploids, 11, 23–25

Hordeum, 122

Inbred families
means, 40, 333–334, 465, 634
relatives, 52
selection, 268, 327–332, 465, 634
variances, 51–54

Inbred lines
means, 40, 333–334, 465, 634
relatives, 52
selection, 268, 327–332, 465, 634
variances, 51–54

Inbreeding
coefficients, 432
estimates of, 442–456
full-sibbing, 180
heterozygosity, 327
selfing, 17, 39
small population size, 436–441
systems, 447
theory, 440

Independent culling, 282–283

Indeterminate gametophyte, 24–25


Interpopulation selection, see Selection

Intrapopulation selection, see Selection

Iodent, 19, 405, 408, 413, 492, 523, 543, 545, 568

Iowa Corn Borer Synthetic 1, 185–186, 610, 612–613, 631

Iowa Ideal, 309–310

Iowa Long Ear, 308

Iowa Stiff Stalk Synthetic
combining ability, 179, 184, 205
correlations, 200
inbreeding depression, 207, 261, 291
testers, 185–191, 197–198
variance estimates, 213

Iowa Two-ear Synthetic, 341, 346, 357, 625

Jala, 402


Jellicorse, 308–310

K

Kitale Composite, 306, 309–310, 316, 318, 632

Kolkmeier, 320–324, 347, 349

Index 659

L
Lancaster Surecrop, 496, 522, 562, 583, 606, 612–613, 630
Line improvement
backcross, 11, 17, 236
convergent, 21
gamete selection, 21–23
haploid, 23–25
pedigree, 17–21, 25, 27
performance, 24, 26
Lines, see Inbred lines
Linkage
bias, 111, 114, 116, 118, 159, 164, 177, 609
disequilibrium, 47, 63, 111, 118, 137, 228, 276–277, 299, 352
effects, 69, 113–114, 149, 157, 164, 170, 642–643

M
Maize
origin, 534–538, 564
races, 542–543, 546
teosinte, 534–537
Tripsacum, 534–537
Mass selection
description, 2
response, 234, 276
theory, 9, 159, 216, 223–285
Mating designs
biparental, 93
cross-classification, 86, 105
design II, 103–113, 160, 180, 613, 622
design III, 113–119, 137–138
epistasis, 142, 159–165
generation means, 206
nested, 76
parent-offspring, 91–96
partial diallel, 133–137
quadrallel, 138–143
special case, 148–159
triallel, 138–143
Maydeae, 534–537
Means
cross between two populations, 42–45, 498
generation, see Generation means
Midland, 185, 298, 330, 332, 345, 347, 405, 418, 462, 488, 521
Modified ear-to-row
description, 245
theory, 245
Monoploids, 11, 22–28

N
Natal Yellow Horsetooth, 341, 613, 633
Nested mating design, 76
Nicotiana, 122
Northern Flints, 2, 546, 548–549, 559

O
Oil, 164, 171–172, 205, 281, 300, 311–314, 346, 353, 383, 448, 569, 647
Origin of maize
Asia, 534–537
classification, 538–542
common ancestor, 72
Corn Belt Dents, 2, 568
teosinte, 534–537
tripartite hypothesis, 535, 537
Ostrinia nubilalis, 328, 360
Outbreeding, 428
Overdominance
gene action, 51, 177, 208, 339, 387, 495, 591
heterosis, 480, 591
inbreeding, 118, 480
selection, 591, 603, 609

P
Parent-offspring regression, 74, 91–96, 150, 152
Paulista Dent, 266, 306, 310, 316, 318, 348, 356
Phenotypic
value, 36, 38, 46, 241, 284, 414
Physoderma maydis, 10, 206
Pioneer Two-ear Synthetic, 185, 346
Piramex, 296, 316, 318, 342–343
Pollen culture, 11, 25
Polytoca, 535
Popcorn, 341, 343, 402, 538–539, 545–546
Populations
composites, 175, 342, 507–510
effective size, 261–262, 267–268, 439–441, 554, 556
exotic, 186, 304–305, 370, 561, 568
genetic variability, 90, 94, 144, 169, 177, 184, 197–198, 225, 261, 330, 355, 507, 554, 560, 631, 636
selection, see Selection
selfing, 157, 271, 500
Prediction
composites, 185, 197, 225
double-cross hybrids, 6, 502, 504
efficiency, 243
formulas, 269, 432, 503, 509–510, 644
hybrid performance, 34, 403
interpopulation, 590
intrapopulation, 12, 337
methods, 6, 209, 462, 499–515
synthetic varieties, 506
theory, 6, 9, 209
three-way cross hybrids, 463, 511
Progeny test, 223, 231, 234, 236, 243, 265, 305, 312, 337, 373, 383, 403, 600, 620
Protaandry, 33, 108, 425
Protein, 24, 164, 205, 312–314, 346, 353, 383, 569
Protogyne, 426
Pure lines, 5, 33, 36, 40, 64, 81, 90–91, 120, 122, 133, 148, 223, 399, 430–431, 470–471, 491, 500, 580
Q
Quadrangle, 138–143, 162, 199
R
Races of maize
classification, 538–542
description, 542–545
distribution, 544–545
European, 545–547
Mexican, 549
multivariate analyses, 542
origin, 534–538
preservation, 542, 552–553
U.S. Corn Belt, 545–552
use, 542–543, 545
Western Hemisphere, 542–545
Random drift, 291, 299, 427
Realized responses, see Selection
Reciprocal full-sib selection, 186, 188–189, 341, 357, 585, 588, 590, 609, 618–629, 633
Reciprocal recurrent selection
half-sib, 611
inbred lines, 395–396, 431, 509, 609, 611, 615, 617
modified, 259, 296
Recurrent selection
interpopulation, 270
intrapopulation, 588
Reid Yellow Dent, 4, 163, 199–201, 203, 304, 315, 488, 496, 522, 543, 548
Rootworms, 361, 587

S
Sclerachne, 535
Score card, 3–4, 403
Selection
backcross, 19–25
combination, 336
correlated response, 276–282, 308, 339, 588
direct response, 215, 303, 330, 337, 339, 347, 454, 585
divergent, 303, 315, 325, 327, 329–330, 336, 410
effective population size, 261–262, 265, 268, 272, 636
efficiency, 234, 236, 283, 291, 373
exotic germplasm, 197, 559–561
expected gain, 238, 243, 247–248, 265, 361, 365, 369, 392–395, 643
fixation of alleles, 276, 443
gamete, 11, 21–23
general effects, 345–358, 622
half-sib, 14–16, 187, 190, 193, 195, 247, 276, 297, 312, 316, 318, 320, 322, 332, 348, 356, 366, 369, 383, 495, 590, 598, 603, 613
among hybrids, 225–228, 384
inbred, 406, 598
inbred testers, 334, 336, 347, 552, 633
increasing gain, 272–276
independent culling, 282–284
indirect response, 303, 327, 337, 351
interpopulation, 590
intrapopulation, 590
linkage effects, 642
measuring changes, 292–299
modified reciprocal recurrent, 259, 296
more than one trait, 282, 550
among populations, 224–225, 631
within populations, 232–233
qualitative traits, 231–236, 305
quantitative traits, 148, 236–255, 305, 328, 373, 647
reciprocal full-sib, 186, 188–189, 341, 357, 585, 588–590, 609, 618–629, 633
reciprocal modified, 15, 259
Selection (cont.)
recombination, 230, 236, 240, 244,
246–251, 254–257, 259, 264, 266–267,
274–277, 311, 325, 360, 366, 372, 427,
436, 440, 560, 563, 582, 586, 589, 591,
598, 600, 602, 625
response, 215, 228, 234, 276, 281,
292–293, 296, 331–332, 334, 340, 351,
436–437, 584
selfed progeny, 16, 335
S1 progeny, 268, 337, 370, 597–600, 633,
636
S2 progeny, 193, 348, 370, 372, 600, 602,
633
tandem, 282–283, 587
tation, 230, 322
theory, 230, 304, 385, 398, 404–409,
412–413, 435–436, 451, 600
Selective advantage, 261
Selfing, 5–6, 14, 17, 20, 26, 28, 39–41, 53–55,
61, 75–76, 78, 91, 104, 133, 152, 157,
180, 182, 236, 241, 248, 251, 262,
270–271, 303, 311, 328, 335, 337, 385,
408, 414, 429, 431–435, 446–447,
449–451, 460, 481, 500, 506–507,
512–513, 519–520, 563, 599, 602, 605,
608, 624, 630, 634, 638
Single-seed descent, 11, 21, 144, 180, 208,
275, 390
Southern Dents, 2, 546, 548–549, 559
Special case (p = q = 0.5), 60–61, 63, 81,
148–159
Specific combining ability, see Combining
ability
Stalk rot, 329, 359–360, 583–585, 587, 589,
596–598, 602, 621
Stiff Stalk Synthetic, see Iowa Stiff Stalk
Synthetic
Synthetic varieties
prediction, 197, 506–507, 512
variability, 175–177, 183
T
Tandem selection, 282–283, 587
Teko Yellow, 341, 343, 613, 633
Teosinte, 534–538, 543
Testcross
estimates, 137–139, 160, 164, 197, 337,
354, 385, 387, 390–391, 414, 457
means, 164, 337, 386, 396
selection, 15, 335, 348, 405–406, 597–598
variance among, 249, 266, 354, 386–389,
395–398
Tester-line interaction, 629
Testers
correlations, 291
general combining ability, 123, 125, 205,
227, 231, 318, 321, 333–334, 383, 440,
466, 518, 600, 609–610, 612, 620, 622,
630
inbred lines, 6, 12–15, 22, 28, 186, 190,
225, 333, 348, 389, 630–631
population improvement, 232, 249, 291,
322, 333, 383–384
specific combining ability, see Combining
ability
stage of testing, 408–415, 603
theory, 333–334
Testing, 5–6, 10, 15–18, 22, 24, 85, 88, 96,
99, 104, 108, 116, 118, 121, 133, 136,
138–139, 144, 146–147, 154, 162, 169,
184, 186, 196–199, 202, 207, 209–210,
215, 228–229, 231, 233, 258, 295,
305–306, 314, 325, 326, 331, 333, 339,
344, 351, 365, 369, 372–373, 383–388,
394, 399, 403–404, 407–415, 417, 431,
440, 456–459, 462–464, 466, 477–478,
486–488, 492, 519–520, 531–532, 562,
568, 590–592, 599–603, 605–606, 608,
610, 612–614, 620–621, 624, 628,
636–637, 644, 648
Tissue culture, 25
Tom Thumb, 402
Triallel, 138–143, 162, 197, 199–200
Trilobachne, 535
Tripsaceae, 534
Tripsacum, 534–537
Triticum, 122
Tuxpeno, 185, 197, 300–301, 549, 559,
561–562, 565–566
U
Ustilago maydis, 300
V
Variance
additive, 9, 46–47, 49–50, 53–54, 56–57,
59, 70, 83, 90, 94, 101, 103, 137–139,
145, 147, 149–150, 153, 157, 162–164,
169–171, 174–179, 196–197, 199–203,
206, 215, 226, 239–240, 246–248,
251–252, 264–265, 268–270, 278, 291,
299, 317, 328, 353–355, 369–370,
392–393, 497, 560, 591, 629, 643–644,
646
cross of two populations, 607
diallel, 125, 134
Index 663

epistasis, 103, 110, 137, 159, 164, 198–210
errors in estimation, 174, 443
inbred families, 39–41, 51–52, 54
inbred populations, 54–55
mating designs, 34, 81–165, 184, 240, 243, 416, 636
minimum variance, 83
negative estimates, 83–84, 162, 200
noninbred families, 51–53
populations, 203
selection experiments, 148, 246
testers, 191
total, 50–51, 173
unbiasedness, 83
use in heritabilities, 57–60, 66, 159, 171–172, 174, 291, 328
use in prediction, 66, 462
Varieties