2006

The quantitative genetics of a non-stiff-stalk maize (Zea mays L) population

Brandon M. Wardyn
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, and the Genetics Commons

Recommended Citation

Wardyn, Brandon M., "The quantitative genetics of a non-stiff-stalk maize (Zea mays L) population" (2006). Retrospective Theses and Dissertations. 1315.
https://lib.dr.iastate.edu/rtd/1315

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The quantitative genetics of a non-stiff stalk maize (Zea mays L.) population

by

Brandon M. Wardyn

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding and Genetics

Program of Study Committee:
Kendall R. Lamkey, Major Professor
Theodore B. Bailey
P. Jeffrey Berger
E. Charles Brummer
Jode Edwards
Jean-Luc Jannink

Iowa State University
Ames, Iowa
2006
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
ii

Graduate College

Iowa State University

This is to certify that the doctoral dissertation of

Brandon M. Wardyn

has met the requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program
TABLE OF CONTENTS

List of Figures iv
List of Tables v
Acknowledgments vi
Abstract vii

Chapter 1. Introduction 1
   References 6

Chapter 2. Literature Review 8
   Inbreeding Depression 8
   Inbreeding and selection 9
   The role of dominance in inbreeding 12
   Quantitative genetic studies 14
   Heritability 17
   References 19

Chapter 3. The Genetic Structure of a Maize Synthetic: the Role of Dominance 26
   Abstract 26
   Introduction 27
   Materials and Methods 31
   Results 40
   Discussion 42
   References 48

Chapter 4. The Relationship Between Genetic Variance Components and Predicted Gains From Selection 60
   Abstract 60
   Introduction 61
   Materials and Methods 65
   Results and Discussion 68
   References 76

Chapter 5. The contribution of Additive and Dominance Effects To Inbreeding Depression 83
   Abstract 83
   Introduction 84
   Materials and Methods 90
   Results and Discussion 95
   References 104
List of Figures

Figure 1. A diagram representing the mating design used for estimation of the genetic covariance parameters. Ovals encompass the generations used in the analysis. 57

Figure 2. Values of the $\sigma^2_D$, $D_1$, and $D_2^*$ expressed as a relative percentage of the respective $\sigma^2_R$ for six environments and the combined analysis for grain yield (A) and plant height (B). 58

Figure 3. Line graphs representing the comparison of observed genetic variances and predicted genetic variances using the estimated genetic covariance parameters for plant height, ear height, grain yield, grain moisture, days to mid pollen (pollen date) and days to mid silk (silk date). Standard error bars represent one standard deviation. 59

Figure 4. The distribution of the change in performance values from F=0 to F=1 for 200 individuals. 110

Figure 5. The distribution of dominance deviations in the S1 and S5 generations. 111

Figure 6. Scatter plots of the dominance deviation and breeding value estimates plotted against the genotypic values for 200 individuals in the S1 generation. 112

Figure 7. Scatter plots of the dominance deviation and breeding value estimates plotted against the genotypic values for 200 individuals in the S5 generation. 113

Figure 8. Scatter plots of the dominance deviation and breeding value estimates for 200 individuals in the S1 and S5 generation. 114
List of Tables

Table 1. Coefficients for genotypic covariance components for the 15 covariance expressions relating five generations of inbreeding. 51

Table 2. Expectations [E(*)], variances [V(*)], and covariances [C(*)] for genotypic values (G), breeding values (A), and dominance deviations (D) of noninbred (F=0) and inbred (F=1) individuals from Edwards and Lamkey (2002). 52

Table 3. Genetic covariance parameters for BSCB1(R) C13 for grain yield, grain moisture, plant height, ear height, days to mid-silk, and days to mid-pollen, for a combined analysis over four locations in 2004 and two locations in 2005. 53

Table 4. Correlations between genotypic values, breeding values, and dominance deviations in both inbred and non-inbred individuals in BSCB1(R)C13. 54

Table 5. Genetic covariance parameters for BSCB1(R)C13 and BS13(S)C0 for grain yield, grain moisture, plant height, ear height, days to mid-silk, and days to mid-pollen, for a combined analysis over six locations. 55

Table 6. Coefficients for genotypic covariance components for the variances of the S0, S1, S3, and S5, generations and the covariance of the S1 generation with the S0, S3, and S5 generations respectively. 79

Table 7. Genetic gain equations for S1, S2, and half-sib selection when the response is measured as the immediate response, the response in the improved population, the response in the improved population of S1 lines, and the response in the improved population of S2 lines. 80

Table 8. Genetic covariance parameters for BSCB1(R)C13 and BS13(S)C0 for grain yield and plant height for a combined analysis over six locations. 81

Table 9. Predicted genetic gain for S1, S2 and half-sib selection methods based on current estimates of the genetic covariance parameters in BSCB1RC(13) and BS13SC(0) for grain yield and plant height for three methods of calculating heritability: the full genetic model, the reduced genetic model, and repeatability. 82

Table 10. Inbreeding depression estimates for five individual environments and combined analysis over all five environments for grain yield. 108
Acknowledgments

I first want to thank my committee members for their insightful suggestions during the course of this research project and their willingness to answer questions whenever asked. Dr. Edwards, you deserve special recognition for the many hours I spent in your office while you laid out genetic theory and statistical procedures in a very effective manner. Your patience and willingness to help did not go unappreciated. I want to thank you Dr. Lamkey, for taking me on as a graduate student and having the confidence in me to allow me to undertake this research project. It has been a very rewarding experience.

I feel very lucky to have worked with such a great group of people at Iowa State, most notably Paul White and the other members of the “corn crew”. The friendships that were created will undoubtedly last longer than my tenure in Ames and I look forward to stopping in whenever my career allows.

I also want to thank my mother and father, Marcia and Leonard Wardyn. You taught me how to look past the trials of everyday life by keeping the long term goals at the center of my thoughts. I could always count on either of you to give me the perspective and encouragement I needed to deal with any challenge.

Finally I would like to thank my wife Michelle. Your dedication and understanding are second to none. The sacrifices you made so I could continue school were many, and I am truly thankful for your support and dedication over the past five years.
Abstract

The genetic relationship among individuals is at the core of nearly all quantitative genetic theory. Dominant gene action has long been either ignored or disregarded as insignificant in many previous genetic models. For grain yield in maize (Zea mays L.), dominance has consistently accounted for a large proportion of genetic variance. We have used previously developed genetic theory that accounts for dominance variance during inbreeding and applied it to a unique breeding design. Our breeding design allowed us to estimate five genetic covariance parameters for six traits. In addition, we developed genetic gain equations that accounted for both dominance and inbreeding. We found that the genetic covariance parameters introduced via inbreeding were significant for five traits. Our estimates of the genetic covariance parameters allowed us to predict genetic gain over a range of selection units and response units. Half-sib selection proved superior to inbred progeny selection when the response was measured in the outbred progeny. In addition, the relative proportions of additive and dominance variance influenced the effectiveness of inbred progeny selection. We also showed that even when dominance constitutes a larger proportion of the total genetic variance than additive variance, the loss of additive effects has a greater influence on the decline associated with inbreeding than the addition of homozygous dominance deviations. Our results also indicated that the reason realized gain often falls short of predicted gain is due to the negative covariance between additive effects and homozygous dominance effects. The effect of a negative covariance is that positive gain via additive effects is offset by negative gain via homozygous dominance deviations.
Chapter 1. Introduction

Plant breeding, like many other disciplines, bases itself on the two most common statistical parameters: the mean and the variance. The goal of most plant breeding programs revolves around shifting the mean in a desired direction via exploitation of the genetic differences between individuals due to the laws of segregation and independent assortment. Without phenotypic variance among individuals, the skills of a plant breeder would be ineffective. Fortunately, phenotypic variance among individuals is common and, depending on the trait, can be quite large. Falconer and Mackay (1996) break down the variance among individuals into a genetic component and an environmental component. The genetic component is of interest due to its repeatability over an infinite set of environments. This genetic component, or stated differently the genetic structure of a population, is determined by allele frequencies and gene effects. Although genetic variance among individuals is a necessity for progress in selection programs, it is not mutually exclusive nor is it the sole indicator of the rate of change that can be made from selection. Regardless of the mode of reproduction, for selection to be effective a covariance between relatives due to genetic effects must exist. The success of selection among individuals thus depends on the narrow sense heritability (Lush, 1936). The covariance between relatives is the numerator in the heritability formula and the phenotypic variance is the denominator. Thus the covariance between relatives as well as the genetic variance will determine the rate of progress in a breeding program.

Hanson (1963) defined heritability as “the fraction of the selection differential expected to be gained when selection is practiced on a defined reference unit”. In a selection
framework, the covariance between relatives is the covariance between the selection unit and the response unit. It is intuitive, that any selection scheme involving plants must have both a clearly defined selection unit as well as a clearly defined response unit. Furthermore, to adequately define the context of a heritability estimate a randomly sampled reference population of genotypes and reference population of environments must be defined (Dudley and Moll, 1969; Nyquist 1991, pp. 239-243). The concept of heritability is not new and has been reviewed and estimated extensively in the literature (Nyquist, 1991), however it is most commonly defined as the “the extent to which phenotypes are determined by genes transmitted from the parents” (Falconer and Mackay pp. 123, 1996). This can be interpreted as the ratio of additive genetic variance to total genetic variance. It should be apparent however, that there is more to a heritability estimate than simply a ratio of two variance estimates.

Of interest to most plant breeders is not the heritability estimate itself, but rather the realized heritability. If we assume a constant amount of genetic variance among the reference population of genotypes, the only ways to increase heritability are to obtain more precise estimates of genotypic values and to maximize the covariance between the selection unit and the response unit. Obtaining precise estimates of genotypic values is a fundamental part of any plant breeding program and thus merits no further discussion. Thus, only the manipulation of the covariance between the selection unit and response unit is at the discretion of the breeder. Historically, breeders have manipulated this covariance via selection methods. Fehr (1991) outlines numerous classical selection methods that have historically been a popular choice among plant breeders. In addition, numerous variations and modifications have been introduced by individuals to fit their budget, time frame, or
biology of the target species. However, to maximize realized heritability it is important to take into account the covariance between the selection unit and the response unit as it is a function of heritability:

\[ h^2 = \frac{\text{Cov}(\text{Selection Unit, Response Unit})}{\sigma^2_{\text{Selection Unit}}} \]

It is intuitive that two factors will determine the covariance between the selection unit and the response unit: the degree of relatedness between selection and response units and the trait under selection.

The resemblance between two relatives is a function of various genetic and environmental sources of variation (Fisher 1918, Wright 1921). As stated previously, the genetic sources are of interest due to the repeatability of genetic effects. Furthermore, the genetic component of this resemblance is a consequence of relatives inheriting copies of the same gene. It is crucial to point out that the degree of relatedness between two individuals can only be defined with respect to a reference population of genotypes. The quantitative genetic theory used to describe the degree of relatedness is built upon the concept of identity by descent. An often overlooked fact is that the probability of two alleles being identical by decent changes when a different reference population of genotypes is considered.

Regardless of the reference population specified, inbreeding will influence the probability of two alleles being identical by descent (F). As this probability increases, the genetic consequences manifest themselves via phenotypic effects at both the population and individual levels. Visualize a population where each individual is the founder of a sub-line derived by self fertilization of one individual. Initially the individuals are non-inbred (F=0) and all of the additive genetic variation is distributed among the individuals. As the
inbreeding process begins, \( F > 0 \) and the total genetic variation is now represented as a mixture distribution of among-line and within-line genetic variance where the mixture distribution is a function of \( F \). When \( F=1 \), the population is composed of completely inbred individuals and all of the genetic variance is among lines. The variance among inbred families is linear in \( F \) under two conditions: i) gene action is strictly additive and ii) gene frequencies equal 0.5 at all loci. When dominance is present or gene frequencies do not equal 0.5, however, the variance among and within inbred families is non-linear in \( F \).

The trait under selection will determine the type of gene action and thus the degree of linearity of \( F \) and variance among inbred families. Given the quantitative nature of most traits, gene frequencies will equal one half only when the population results from a two parent cross. Although two parent populations are utilized by breeders, when considering natural populations, two parent populations are rare. Thus, a plant breeder working with a natural population will never be operating under the conditions necessary to assume a linear relationship between \( F \) and the variance among inbred families. We can extend this non-linearity to include the variance within inbred families as well using the same reasoning. It follows that the covariance between individuals within the same sub-line will also be non-linear in \( F \).

This non-linear covariance between related individuals is the focal point of this dissertation. When planning a breeding program, the response per year is an important factor in the final decision. As outlined earlier, a component to the response is the covariance between two related individuals. This leaves us with the problem of modeling this non-linear covariance between two related individuals. Although seldom recognized, this covariance can be modeled via population genetic covariance parameters (Harris 1964, Gillois 1964). If
these population genetic covariance parameters are known, a breeder can make accurate predictions as to the response from a given selection program and thus increase their probability of success.
References

Hanson, W. D. and H. F. Robinson (ed.) Statistical Genetics and Plant Breeding,

Dudley, J.W., and R.H. Moll. 1969. Interpretation and use of estimates of heritability and
genetic variances in plant breeding. Crop Sci. 9:257-262.

Group Limited, London


Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian

Springer-Verlag, Germany.

Hanson, W. D., 1963. Heritability. p. 125-140. In: Statistical Genetics and Plant Breeding,
Washington, D. C.


Lush, J. L., 1936. Genetic aspects of the Danish system of progeny-testing swine, Iowa

Chapter 2. Literature Review

Inbreeding Depression

Inbreeding is the consequence of the mating between two related individuals. The degree of inbreeding is quantified by the parameter $F$ which represents the probability that two alleles are identical by descent. The two alleles in question can be contained within an individual or among individuals. Inbreeding plays an important role in plant breeding in two ways. The first being that a completely inbred individual will breed true and its genetic effects can be predicted. The second role of inbreeding in plant breeding programs is the manifestation of inbreeding depression.

Inbreeding depression can be defined as a decline in fitness due the increased probability of alleles being identical by descent. Hallauer and Miranda (1988 pp. 314-315) outline inbreeding depression estimates in maize for eight studies (Sing et al., 1967; Genter, 1971; Harris et al., 1972; Hallauer and Sears, 1973; Cornelius and Dudley, 1974; Good and Hallauer, 1977; Levings et al., 1967; Rice and Dudley, 1974) which vary in their method of inbreeding and germplasm evaluated. The results were consistent across all studies in that inbreeding depression was observed for all grain and plant traits with the lone exception being days to anthesis. The results presented by Hallauer and Miranda (1988) agree with the findings of San Vicente and Hallauer (1993) who analyzed pre-1960 and post-1970 inbred lines rather than closed populations. Good and Hallauer (1977) compared three methods of inbreeding (selfing, full-sibbing, and full-sibbing followed by selfing) and found no differences between the methods of inbreeding for plant height which agreed with the results
reported by Cornelius and Dudley (1974). In terms of yield, the only significant difference between the methods was at near complete homozygosity where selfing and full-sibbing followed by selfing were significantly different. Benson and Hallauer (1994) found that the rate of inbreeding depression decreased for yield in both BSSS and BSCB1 populations after nine cycles of selection. Similar to previous studies they also found a reduction in the mean associated with inbreeding for all traits except days to anthesis.

Keeratinijakal and Lamkey (1993:73-77) calculated inbreeding depression in absolute units as the difference between the mean of population per se and the mean of the S1 lines for BSSS(R), BSCB1(R), and BSSS(R) X BSCB1(R). They found a general increase in inbreeding depression for the population cross from 1.01 to 2.32 Mg ha\(^{-1}\) from cycle 0 to cycle 11. No general trend was detected for BSCB1(R) whereas BSSS(R) showed a decrease in inbreeding depression from 1.15 Mg ha\(^{-1}\) in cycle 0 to 0.64 Mg ha\(^{-1}\) in cycle 11. They attributed the increase in inbreeding depression in the population cross to selection at complimentary loci in each population.

**Inbreeding and Selection**

A major concern of recurrent selection programs is a loss of genetic variance due to finite population sizes. Small effective population sizes lead to a reduction in genetic variance via a loss of alleles and is referred to as genetic drift. Genetic drift is nothing more than inbreeding at the population level. Guzman and Lamkey (2000) investigated the effect of varying the effective population size for five cycles of selection in the maize population BS11. Five cycles of S1-progeny recurrent selection were carried out in BS11 under 4
effective population sizes: 5, 10, 20, and 30. The expected level of inbreeding after five cycles of selection ranged from 0.38 in the 5-S1 program to 0.08 in the 30-S1 program. They reported no significant difference in the amount of additive genetic variance among the selection schemes. According to Crow and Kimura (1970) the expected level of inbreeding in the 5-S1 program should have decreased genetic variance drastically. Guzman and Lamkey (2000) hypothesized that the lack of reduction in additive genetic variance was due to a low frequency of favorable alleles or the conversion of non-additive genetic variance (i.e. dominance and epistasis) to additive genetic variance. Even though the genetic variance in each of the BS11 populations was unaffected by population size, response to selection for yield was affected. Weyhrich et al. (1998) compared the yield of the 5-S1, 10-S1, 20-S1 and 30-S1 BS11 populations and found that grain yield increased in all populations except for the 5-S1 population. They found a decrease of 0.22 Mg ha-1 per cycle for the 5-S1 population. Thus it appears that although no significant loss of genetic variance was observed with small population sizes in BS11, response to selection was significantly affected.

The results put forth by Guzman and Lamkey (2000) and Weyhrich et al. (1998) in BS11 were after progeny selection. Walters et al. (1991) used BSSS(R), which was developed via reciprocal recurrent selection (Comstock et al., 1949), as a germplasm source and found a significant increase in grain yield and a significant decrease in moisture after nine cycles. They noted a general decrease in genetic variance among S1 lines for four traits after nine cycles of selection although not all decreases were statistically significant. This was in agreement with the findings of Lantin and Hallauer (1981) who reported that genetic variability had not decreased after four cycles of selection in the reciprocal full-sib recurrent selection program between BS10 and BS11. Hallauer (1984) also evaluated the BS10-BS11
reciprocal full-sib program and found no reduction in genetic variance for yield after seven cycles of selection. However, it is interesting to note that Frank and Hallauer (1999) found that after 10 cycles of reciprocal full-sib selection in BS10-BS11, significant reductions in genetic variance were found for grain moisture, root lodging, stalk lodging, and days after planting to mid-silk. No reductions in genetic variance were found for grain yield, plant height, and ear height.

Thus it appears that a large body of evidence exists to support the hypothesis that, in terms of recurrent selection, selection does not significantly reduce genetic variance for grain yield. Thus it is logical to test a hypothesis which asks if genetic drift is a concern in closed population/recurrent selection programs. Keeratinijakal and Lamkey (1993:78-82) performed an experiment where they evaluated cycles 0, 4, 7, 9, and 11 of both BSSS(R) and BSCB1(R) as well as cycles eight and 10 of the interpopulation cross BSSS(R) X BSCB1(R). They used the model described by Smith (1983) to estimate genetic parameters. Under Smith's model, the parameter DQI represents the effect of loss of heterozygotes, or stated differently, the change in performance due to inbreeding. According to their results, both populations per se experienced a net loss of 0.012 Mg ha-1 per cycle due to the loss of heterozygotes. Furthermore, they reported that favorable alleles with additive effects as well as favorable alleles with dominance effects both increased in frequency in the two populations. However, the response of the populations per se for grain yield was not improved which supports the hypothesis that inbreeding depression is resulting from genetic drift. Several studies (Tanner and Smith, 1987; Helms et al., 1989; Eyherabide and Hallauer, 1991,) previously reported results in agreement with Keeratinijakal and Lamkey's findings.
The Role of Dominance in Inbreeding

Given the overwhelming evidence that genetic variance does not necessarily decline with recurrent selection and that genetic drift is in fact impacting selection response via inbreeding depression, it is clear that additive genetic theory put forth by Wright (1951) is not adequate to describe the genetic architecture of a maize population under selection. It has been shown theoretically that with the presence of dominant or epistatic gene action, inbreeding does not necessarily reduce genetic variance (Robertson, 1952; Avery and Hill, 1979; Bryant et al., 1986a,b; Goodnight, 1987, 1988, 1995; Cockerham and Tachida, 1988; Tachida and Cockerham, 1989; Whitlock et al., 1993; Willis and Orr, 1993; Cheverud and Routman, 1995, 1996; Wang et al., 1998). Edwards and Lamkey (2003) used empirical estimates obtained in BS13 to predict changes in genetic variance due to population subdivision and subsequent accumulation of inbreeding. Their predictions showed that additive genetic variance within a subpopulation would initially increase up to an Fst value of approximately 0.4. Further inbreeding beyond an Fst of 0.4 would result in a loss of additive genetic variance within the subpopulation. They also found that the predicted genetic variance among non-inbred individuals was greater than the additive expectation for all traits. It is also clear from the work of Edwards and Lamkey (2003) that dominant gene action has a large impact on not only total genetic variance but also additive variance.

The degree of dominance in maize has been fiercely debated in the past with proponents falling into two camps: overdominance and dominance. In an F2 population, the degree of dominance is usually greater than one (Robinson et al., 1949; Gardner et al., 1953; Gardner and Lonnquist, 1959; Moll et al., 1964; Han and Hallauer, 1989) which corresponds
to overdominance. However, random mating the F2 population usually reduces the degree of dominance to less than one (Gardner and Lonnquist, 1959; Moll et al., 1964; Han and Hallauer, 1989) which corresponds to partial or complete dominance. The degree of dominance is of utmost importance to plant breeding programs as it will determine the effect that inbreeding will have on genetic variance. However, what is of interest to this study is the influence of dominance on inbreeding depression and the variance of inbreeding depression within a population. It has been fairly well recognized in evolutionary biology that the variance of inbreeding depression has a large influence on mating system evolution (reviewed in Kelly, 2004), however work in crops species is relatively non-existent.

The variance of inbreeding depression among lines developed from the same population is of great interest to maize breeders. To the authors’ knowledge, no studies have investigated the cause of differential rates of inbreeding depression among maize lines developed from the same population. Numerous studies have identified rates of inbreeding depression at the population level, but failed to address inbreeding depression at the level of the individual. In a selection program, selection is practiced on individuals and thus we argue that inbreeding depression defined at the level of the individual is the correct, due to its utility, definition of inbreeding depression. Thus, what would be most useful to breeders is a quantification of the genetic effects of inbreeding at an individual level. Falconer and Mackay (1996) define the genotype of an individual as the sum of an individual’s breeding value and dominance deviation. They define the breeding value as “twice the mean deviation of the progeny from the population mean” (Falconer and Mackay p.114, 1996). Breeding value can also be defined as the summation of additive effects. Fisher (1918) defines additive effects via a regression of genotypic values on actual genotypic frequencies. Harris
(1964) extended Fisher’s genetic model to include inbred relatives and defines additive effects with reference to a panmictic reference population. As discussed in Edwards (2006), under Fisher’s model, additive effects can change with inbreeding whereas in Harris’s model additive effects never change. This is an important development made by Harris (1964) that has not been readily used in empirical studies and not explored theoretically until Edwards (2006).

The beauty of Harris’s development was that now we can define additive effects for multiple generations of inbreeding and thus also obtain breeding values for inbred generations. Following Falconer and Mackay’s definition of genotype, the difference between an individual’s genotypic value and breeding value is the dominance deviation which is independent of an individual’s inbreeding coefficient. Thus the measure of inbreeding depression for an inbred individual is simply the homozygous dominance deviation. It follows that these homozygous dominance deviations have an expected value and a variance, which are both dependent on F (Edwards and Lamkey, 2002). Not surprisingly, there is also a covariance between genotypic values, breeding values, and dominance deviations which are all dependent on F. Thus it is possible to obtain individual inbreeding depression and breeding value estimates and at the same time quantify these effects at the population level, although it has rarely been done in the literature.

Quantitative Genetic Studies

Several studies do exist that have investigated the quantitative genetic properties of inbreeding. Coors (1988) evaluated the response to half-sib and S1 recurrent selection in a
narrow based maize synthetic. He found a significant reduction in inbreeding depression after four cycles of selection following a combined half-sib and S1 selection procedure (for details concerning the procedure see Goulas and Lonnquist, 1976). Coors also found a large negative covariance between additive effects and inbred homozygous dominance deviations (D1). These were the first published estimates of D1 in any species. Coors hypothesized that the lack of effectiveness of inbred progeny selection was due to D1 affecting the variance among inbred progenies.

Shaw et al. (1998) investigated the genetic components of flowering time and morphology in a *Nemophila menziesii* population that was undergoing inbreeding. The model they used was the same model developed by Harris (1964) and refined by Cockerham (1971) and Jiang and Cockerham (1990). Their breeding design consisted of three generations that contained a wide array of inbreeding coefficients and genetic relationships. They found strong inbreeding depression and a significant impact of additive variance. What is of interest to this paper is that they found that D2* (the variance of homozygous dominance deviations) contributed significantly to petal length, petal width, and flowering date (all reproductive traits). However, they found no influence of the inbreeding components for plant size measurements. Another interesting result of their study was that they found significant inbreeding depression for the size traits, but failed to identify any variance of homozygous dominance deviations. They suspect that a possible reason for this is that many loci are contributing to inbreeding depression and thus the individual effect at a locus is small and thus the variance cannot be detected.

Abney et al. (2000) were the first to conduct an empirical study concerning this topic in humans. They analyzed a data set of the Hutterite population for levels of HDL (the
favorable type of cholesterol). Numerous models were tested with the data set and they also ran simulations varying the average inbreeding level, level of dominance, and sample size. They found insignificant inbreeding depression in all of the models they tested with the real data set. Thus it is not surprising that they found no significant inbreeding dominance components. The simulation portion of their study indicated that a fully dominant model gave them the most power to detect inbreeding dominance components. It was also apparent in their results that the sample sizes they used (1,000+), gave them enough power to discern if estimates were different from zero, but not enough power to obtain accurate estimates of the inbreeding dominance components. They concluded that in terms of genetic mapping, the driving force behind their study, inbreeding dominance components could be left out of the model when modeling for background effects of HDL in this population.

Edwards and Lamkey (2002) analyzed the maize population BS13 and used the genetic model of Harris (1964). They calculated probabilities of IBD by Cockerham’s suggestions (1971, 1983). The breeding design used took into account suggestions made previously (Cornelius and Van Sanford, 1988 and Cockerham, 1983) and included inbred progenies from early in the inbreeding process in addition to outbred progenies of the inbred material. Significant inbreeding depression was found for all of the traits investigated. They found, in general, a significant contribution of all the inbred variance components to genetic variance. In their analysis, they found that the variance of inbred dominance deviations was 2.65 times the variance of noninbred dominance deviations. They also found a negative correlation between inbred dominance deviations and breeding values in the BS13 population.
The idea that inbreeding depression is a variable and selectable trait was put forth by Pray and Goodnight (1995). The large variability of inbred dominance deviations found by Edwards and Lamkey (2002) supports this idea. The correlation between inbred dominance deviations and genotypic values is the degree to which selection is acting on inbreeding depression. Stated differently, if the correlation between inbred dominance deviations and genotypic values is high, we will actually be selecting for reduced inbreeding depression. This would explain the reduced effectiveness of inbred progeny selection in maize and agrees with the results found by Coors (1988). As hypothesized by Edwards and Lamkey (2002), selection on inbred progenies is selecting for genotypic values and individuals with reduced inbreeding depression will have higher inbred progeny means. However, the higher inbred progeny means correspond to lower outbred progeny means due to $D_1$ being negative.

**Heritability**

The parameter $D_1$ is not well researched in that few empirical estimates are available. Furthermore, the concept of dominance influencing genetic variance and thus heritability is usually overlooked as dominance is often assumed to be zero or negligible. It has been repeatedly shown in maize that dominance is important for grain yield but what is not addressed is the influence of dominance on heritability. Hanson (1963) defines heritability as “the fraction of the selection differential expected to be gained when selection is practiced on a defined reference unit”. In a mathematical sense, the heritability is the regression coefficient of the regression of the response unit on the selection unit. If we decompose the
regression coefficient we can see that it is the covariance between the genotypic values of the selection and response units divided by the genetic variance of the selection unit:

\[ h^2 = \frac{\text{Cov}(G_{\text{selection unit}}, G_{\text{response unit}})}{\sigma^2_{\text{selection unit}}}. \]

Jacquard (1983) coins this representation of heritability as the biometric heritability.

A common approach to estimating heritability in the literature is to regress offspring values on parental values. Under an ideal additive model, this approach is correct. However, when the additive model assumptions are violated, the parent-offspring regression is most often biased upward. As alluded to previously, the cause for this bias in most maize breeding programs is most often the genetic covariance parameters introduced by the combination of inbreeding and dominance. Cockerham (1983) outlines the model that accurately estimates the covariance between two individuals regardless of the level of dominance or inbreeding level. Using Cockerham's (1983) notation, inbreeding and dominance introduce three new parameters not customarily used:

- \( D_1 \), the covariance between additive effects and inbred dominance effects,
- \( D_{2*} \), the variance of homozygous dominance deviations,
- and \( H^* \), the square of inbreeding depression.

Holland et al. (2003) discusses some of the implications of these three parameters on heritability estimation. The effects of \( D_1 \), \( D_{2*} \) and \( H^* \) on heritability are directly related to the amount of inbreeding and dominance in a population. The magnitude of the effect of the three parameters is determined by the amount of inbreeding and the amount of dominance.
References


Chapter 3. The Genetic Structure of a Maize Synthetic: the Role of Dominance

A paper to be submitted to Crop Science.

Brandon M. Wardyn, Jode W. Edwards and Kendall R. Lamkey

Abstract

In selection programs, the covariance between parents and offspring largely determines the success of selection. We have estimated the variances and covariances between noninbred individuals and both their inbred and outbred progeny in the non-stiff stalk maize population BSCB1(R)C13. Estimation of these variances and covariances has allowed us to estimate the genetic covariance parameters for BSCB1(R)C13. Previous estimates of genetic covariance parameters in maize have been used to describe the ineffectiveness of inbred progeny selection in the stiff stalk population BS13. Our estimates in BSCB1(R)C13 indicated that the dominance variance was larger than the additive variance for grain yield whereas the additive variance was larger than the dominance variance for all other traits. Negative estimates of the covariance between additive and homozygous dominance deviations were found for all traits with the exception of traits associated with reproductive maturity, suggesting a negative relationship between inbred and outbred performance. A direct result of this genetic variance structure was the correlation between genotypic values and breeding values was lower for grain yield than any other trait. Our results were similar to previous results found in the stiff stalk maize population BS13. Thus it appears that similar genetic variance structures have been formed by selection in two maize populations that differ in their genetic background.
Introduction

The debate about the predominant type of dominance in maize, dominance or overdominance, has been prevalent in the maize (*Zea mays* L.) community for a number of years (Crow, 2000). What has never been a major source of debate, however, is that dominance exists and is important in maize. Hallauer and Miranda (1988) reported that over an average of 99 studies, the dominance variance for grain yield was 286.8 g plant$^{-1}$ whereas the estimate for the additive variance was 469.1 g plant$^{-1}$. They also reported an average dominance to additive ratio of 0.94 for grain yield and 0.53 for plant height. Thus at least in terms of grain yield, dominance variance constitutes a large portion of the total genetic variance. Grain yield is unique in maize as it routinely shows a higher proportion of dominance variance than other traits.

The presence of dominance, specifically directional dominance, coupled with inbreeding can have a large impact on both individual and population performance via inbreeding depression. Hallauer and Miranda (1988 pp. 314-315) outline inbreeding depression estimates in maize for eight studies (Sing et al., 1967; Genter, 1971; Harris et al., 1972; Hallauer and Sears, 1973; Cornelius and Dudley, 1974; Good and Hallauer, 1977; Levings et al., 1967; Rice and Dudley, 1974) which vary in their method of inbreeding and germplasm evaluated. The results were consistent across all studies in that inbreeding depression was observed for all grain and plant traits with the lone exception being days to anthesis. Numerous other studies have been performed in maize which measure inbreeding depression at the population level, but to the authors’ knowledge no studies have estimated the variability in inbreeding depression among individuals within the same maize population.

Several studies exist which have investigated the quantitative genetic properties of
inbreeding with dominance while using Harris’s (1964) model. Coors (1988) evaluated the response to half-sib and S1 recurrent selection in a narrow based maize synthetic. He found a significant reduction in inbreeding depression after four cycles of selection following a combined half-sib and S1 selection procedure. Coors (1988) also found a large negative covariance between additive effects and homozygous dominance deviations ($D_1$). These were the first published estimates of $D_1$ in any species. Coors hypothesized that the lack of effectiveness of inbred progeny selection was due to $D_1$ affecting the variance among inbred progenies. Shaw et al. (1998) investigated the genetic components of flowering time and morphology in a *Nemophila menzieissi* population that was undergoing inbreeding. Their breeding design consisted of three generations that contained a wide array of inbreeding coefficients and genetic relationships. They found that $D^*_2$ (the variance of homozygous dominance deviations) contributed significantly to petal length, petal width, and flowering date (all reproductive traits). However, they found no influence of the inbreeding components for plant size measurements. Another interesting result of their study was that they found significant inbreeding depression for the size traits, but failed to identify any variance of homozygous dominance deviations.

Edwards and Lamkey (2002) analyzed the maize population BS13 with the genetic model of Harris (1964). Probabilities of identity by decent were calculated by Cockerham’s suggestions (1971, 1983). Their breeding design took into account suggestions made previously (Cornelius and Van Sanford, 1988 and Cockerham, 1983) and included inbred progenies from early in the inbreeding process in addition to outbred progenies of the inbred material. Significant inbreeding depression was found for all of the traits investigated. They found, in general, a significant contribution of all the inbred variance components to genetic...
variance. In their analysis, they found that the variance of inbred dominance deviations was 2.65 times the variance of noninbred dominance deviations. They also found a negative correlation between inbred dominance deviations and breeding values in the BS13 population.

The idea that inbreeding depression is a variable and selectable trait was put forth by Pray and Goodnight (1995). The large variability of inbred dominance deviations found by Edwards and Lamkey (2002) supports this idea. The correlation between genotypic values of inbred individuals and inbred dominance deviations is a measure of the correlation between the value of an inbred individual and inbreeding depression because the expectation of inbred dominance deviations, namely least-squares homozygous dominance deviations $\delta_i$, are the quantitative genetic basis for inbreeding depression. The correlation between inbred dominance deviations and breeding values is an estimate of the correlation between inbred performance and outbred performance. For grain yield, Edwards and Lamkey (2002) found a correlation between inbred genotypic value and inbred dominance deviations of 0.63 and between inbred genotypic value and breeding value of 0.34, suggesting that selection among inbred individuals in BS13(S)C0 would have a greater impact on changing inbreeding depression than it would on changing inbreeding outbred performance. Given the importance of inbreeding in maize programs, a better understanding of how inbreeding affects genetic values of individuals and how inbreeding affects covariances of relatives is needed, and in more maize populations than have been studied. A better understanding of inbreeding may help to uncouple the inbred and hybrid performance of maize lines, which is of great interest to the commercial maize industry. The major questions being asked by this study include I) What is the importance of dominance in a maize synthetic? II) How does
inbreeding affect the genetic variance among individuals in BSCB1(R)C13? III) How do the
genetic covariance parameters in BSCB1(R)C13 compare to BS13? IV) How does the
current breeding design compare to previous designs used to estimate the genetic covariance
parameters?
Materials and Methods

Germplasm

The maize (*Zea mays* L.) population Iowa Corn Borer Synthetic No. 1 (BSCB1), a member of the non stiff stalk heterotic pattern, was the source of germplasm in this study. BSCB1 was developed in the 1940’s at Iowa State University via an intermating of 12 inbred lines (Penny and Eberhart, 1971). See Hagdorn et al. (2003) for an outline of the 12 parents. BSCB1 was developed via reciprocal recurrent selection (Comstock et al., 1949) with Iowa Stiff Stalk Synthetic. Details for the first five cycles of selection in BSCB1(R) can be found in Penny and Eberhart (1971). Please refer to Holthaus and Lamkey (1995), Keeratinijakal and Lamkey (1993), and Schnicker and Lamkey (1993) for details concerning the selection and breeding methods for cycles six through 11. The same procedure used in cycle 11 was followed for cycles 12 and 13. For the current study, cycle 13 of BSCB1 [BSCB1(R)C13] was the germplasm source.

Mating Design

Four hundred random S0 plants were self-pollinated in the BSCB1(R)C13 population. Resulting S1 ears were planted ear-to-row and the first three plants in each row were self-pollinated. One ear was randomly chosen and the resulting S2 ear was planted the following year and the first three plants were self-pollinated. This process was repeated until S5 seed from the initial 400 S0 plants was obtained. Thus, each S0 line was represented in five generations of inbreeding where each generation was a direct descendant of the initial S0 plant (Fig. 1). The only selection applied during the inbreeding process was that enough seed be present to plant a full nursery row the following year. In the summer of 2003, the S1, S3, and S5 generations were planted in a nursery row 3.8 m long at a density of 15 plants per
row. Seed quantities were increased via sib-mating within a nursery row. Effort was made
to use each plant once as either a male or a female and reciprocal crosses were not made. A
balanced bulk of approximately 10 ears from each nursery row was made and used as the
source for yield trial plots. Due to poor stands, some rows were re-grown and subsequently
sib-mated in the 2003-2004 winter nursery in order to obtain adequate quantities of seed for
yield trials.

In addition to nursery rows, each S1 and S5 line was planted in isolation with
BSCB1(R)C13. The S1 and S5 lines were detasseled and used as females, being open-
pollinated with BSCB1(R)C13 used as the male (Fig. 1). The isolation rows were 5.49 m
long and were planted at a density of 20 plants per row. All plants in a row were harvested
and shelled in bulk. Harvested seed was treated with Maxim XL (Syngenta Crop Protection
Inc. Greensboro, North Carolina) at the labeled rate for use in replicated yield trials.

Experimental Design

The 200 lines in five generations of inbreeding were planted near Ames, Carroll,
Crawfordsville, and Rippey, IA, in 2004 and near Ames and Carroll, IA in 2005. Each
location contained two replications of the experiment, laid out in a split-plot design with
generation of inbreeding (S1, S3, S5, S1 topcrossed, S5 topcrossed) as the whole plot factor
and individual lines within generations as subplots. Whole plots were laid out in a
randomized complete block design. Sub-plots were randomized in a 10 by 20 row-column
alpha lattice \([a(0,1)]\) within each generation of inbreeding. The 200 entries were randomly
selected from approximately 350 entries with an adequate amount of selfed seed to plant the
necessary nursery and isolation rows for seed increase. Bulks were made of each generation
and used as border rows that surrounded each whole plot. An individual yield trial plot
consisted of two rows, both 5.49 m in length with 0.76 m between rows. Seeds were machine planted at a density of 76 540 plants ha\(^{-1}\) and thinned to 62 190 plants ha\(^{-1}\). Again, cultural practices were consistent of commercial maize production in central Iowa.

Data were collected on an individual plot basis for days to mid-silk (days after planting), days to mid-pollen shed (days after planting), root lodging (%), stalk lodging (%), harvestable grain weight (g adjusted to 15.5 % grain moisture), and grain moisture (%). Days to mid-silk were determined when half of the plants in a plot had visible silks and days to mid-pollen were determined when half of the plants in a plot had begun shedding pollen. Plant height (cm) and ear height (cm) were obtained by measuring 10 plants per plot on the out crossed generations and six plants per plot for the S1, S3, and S5 generations. Plant height was the distance between the soil surface and the uppermost leaf collar. Ear height was measured as the distance between the soil surface and the uppermost ear node. The mean value for plant height and ear height was calculated on a per plot basis and used in the analysis. A reduced number of plants were measured in the inbred generations due to the lack of variability within plots for inbred generations. Grain yield was on a harvestable weight basis as all plots were machine harvested.

**Genetic Model**

The same genetic model used by Edwards and Lamkey (2002) was used for this study. The model is based on an extension of Fisher’s (1918) genetic model by Harris (1964) to include inbred relatives. Harris developed a completely general parameterization of the genetic covariance between two individuals with arbitrary levels of inbreeding. The general parameterization requires the following assumptions i) there is no linkage among the loci that influence the traits being evaluated, ii) the original population is random mating, iii) there has
been no selection practiced during the development of the two individuals, and iv) the
individuals have autosomal diploid loci.

Harris (1964) defines the genetic value of an individual:

\[ g_{ij} = \mu + \alpha_i + \alpha_j + \delta_{ij} \]

where:

- \( g_{ij} \) = the expected phenotypic value of an individual with genotype \( A_iA_j \)
- \( \mu \) = the population mean
- \( \alpha_i \) = the additive effect of allele \( A_i \)
- \( \delta_{ij} \) = the dominance deviation of genotype \( A_iA_j \)

Under this model, the covariance between two individuals (X and Y) can be represented by:

\[
\text{Cov}(X,Y) = 2\theta_{XY}\sigma_A^2 + 2(\Delta_{\bar{X}Y} - \delta_{\bar{X}Y})\sigma_D^2 + 2(\gamma_{\bar{X}Y} + \gamma_{XY})D_1 + \delta_{\bar{X}Y}D_2^* + (\Delta_{\bar{X}Y} - F_{\bar{X}Y})H^* 
\]

where:

- \( \sigma_A^2 \) = additive genetic variance
- \( \sigma_D^2 \) = dominance variance
- \( D_1 \) = covariance between additive effects and homozygous dominance effects
- \( D_2^* \) = the variance of homozygous dominance deviations
- \( H^* \) = the sum of homozygous dominance deviations, squared, and
- \( \theta_{XY}, \Delta_{\bar{X}Y}, \delta_{\bar{X}Y}, \gamma_{\bar{X}Y}, \gamma_{XY}, \Delta_{\bar{X}Y}, F_{\bar{X}Y}, \) and \( F_{\bar{Y}} \) are probabilities of identity by descent for 2, 3, or 4 alleles.

(collectively referred to as genetic covariance parameters).

Cockerham (1971) describes the calculation of the 15 probabilities of identity by decent for
two, three, and four alleles which can be reduced to the eight descent measures in the expression of covariances between inbred relatives. The eight probabilities determine the coefficients for each of the five genetic covariance parameters \( \sigma^2_A, \sigma^2_D, D_1, D_2^*, H^* \).

For our study, we had five generations of inbreeding within each independent line and thus 10 covariances among the generations and 5 variances. The coefficients used for the genetic covariance parameters for each of the 15 (co)variances are listed in Table 1. The coefficients were calculated based on the reports of Cockerham (1971, 1983) as well as Cockerham and Matzinger (1985).

**Influence of Inbreeding**

It should be emphasized that under the Harris (1964) model, inbreeding does not change any of the population specific genetic covariance parameters, but it does change the relative contributions of the parameters to genotypic (co)variances. To outline the changes induced by inbreeding it is necessary to outline the genetic makeup of an individual. Falconer and Mackay (1996) define the genetic value of an individual as:

\[
G = A + D
\]

where:

- \( G \) = the genotypic value of an individual
- \( A \) = the breeding value of an individual
- \( D \) = the difference between the genotypic value and the breeding value of an individual.

\( G, A, \) and \( D \) from the Falconer and Mackay model are estimators of the actual genetic effects in the Harris model. \( G, A, \) and \( D \) are observed values and their expected values change in accordance to the reference population.

Following the Harris (1964) model, the values of \( G, A, \) and \( D \) are consistently defined
with respect to a panmictic reference population and are thus independent of inbreeding. Inbreeding changes the expected values, variances, and covariances between G, A, and D (Table 2). In addition, inbreeding introduces additional genetic covariance parameters into the expected values, variances, and covariances where the degree of change between a noninbred individual and an inbred individual is determined largely by $D_1$, $D_2^*$, and $H^*$. The utility of Harris’s model (1964) is the ability to describe the covariance between any two individuals regardless of their respective F values, which is not possible with the Fisherian model.

**Data Analysis**

Grain yield, grain moisture, plant height, ear height, days to mid-silk, and days to mid-pollen were the traits analyzed. All traits were analyzed using a mixed linear model where environments (location-year combinations) and replications within environments were considered fixed effects. Environments were considered fixed as they were not of primary interest in this study. For grain yield and grain moisture, lattice rows and columns were fitted as fixed effects. Only lattice rows were fitted for plant height and ear height. The random effect vector contained only the line effect. Here a line takes the form $u'_{ijk}$ where:

- $i = \text{the } i^{th} \text{ line } (i = 1 \ldots 200)$
- $j = \text{the } k^{th} \text{ environment } (k = 1 \ldots 6)$
- $k = \text{the } j^{th} \text{ generation } (j = 1 \ldots 5)$

such that $u'(25,3,1)$ represents the 25th line at the third environment in the first generation. In this model, lines are cross classified with environments and generations.

The 200 lines were considered independent in the model because each line was developed from an independent S0 individual in the founder population. Thus all lines share
the same variance-covariance matrix because they had the same pedigree and different lines
grown in the same or different environments have a covariance of zero. Given six
environments and five generations, the variance-covariance matrix takes the form of a 30 by
30 matrix. Due to the genetic design being used, the 30 by 30 variance-covariance matrix is
a linear function of the genetic covariance parameters and genotype by environment
interaction. Representing the vector of five random effects for five generations of inbreeding
of line \( i \) observed in environment \( j \) as \( u_{i,j} \), the covariance between a genetic effects of a line
grown in environments \( j \) and \( j' \) can be expressed as:

\[
\text{Cov}(u_{i,j}, u_{i,j'}) = A_1 \sigma_A^2 + A_2 \sigma_D^2 + A_3 D_1 + A_4 D_2^* + A_5 H^*.
\]

whereas the variance of a line grown in the same environment can be expressed as:

\[
\text{V}(u_{i,j}) = A_1 \sigma_A^2 + A_2 \sigma_D^2 + A_3 D_1 + A_4 D_2^* + A_5 H^* + A_1 \sigma_{AE}^2 + A_2 \sigma_{DE}^2 + A_3 D_{1E} + A_4 D_{2E}^* + A_5 H_{E}^*.
\]

The matrices \( A_1, A_2, A_3, A_4, \) and \( A_5 \) are 5 by 5 coefficient matrices that are composed of the
coefficients listed in Table 1. These matrices describe the expected genotypic variance and
covariances among the five generations. The parameters \( \sigma_{AE}^2, \sigma_{DE}^2, D_{1E}, D_{2E}^* \) and \( H_{E}^* \)
represent the common environmental effect shared by genotypes grown in the same
environment. They are equivalent to genotype by environment interaction effects in
traditional analysis of variance models.

All effects were fit in a mixed linear model in SAS Version 9 via the mixed
procedure. From this, Restricted Maximum Likelihood (REML) estimators of the ten
genotypic covariances were obtained. Because the exact sampling distribution of variance
components is unknown, we relied on asymptotic large-sample variances and covariances of
the genetic covariance parameters obtained directly from the mixed procedure. From this, standard errors were attached to each estimate and significance was assigned at the 0.05 level if the estimate was more than two standard errors away from zero. This asymptotic variance covariance matrix was also used to calculate the correlations among the genetic covariance parameters. Error variances were found to be heterogeneous by generation and were fit accordingly in the analysis.

Correlations between G, A, and D for inbred individuals were calculated as:

\[ r_{G,A} = \frac{2(\sigma_A^2 + D)}{\sqrt{2\sigma_A^2 (2\sigma_A^2 + 4D_1 + D^*_1)}} \]

\[ r_{G,D} = \frac{2D_1 + D^*_1}{\sqrt{D_1^* (2\sigma_A^2 + 4D_1 + D^*_1)}} \]

and

\[ r_{A,D} = \frac{2D_1}{\sqrt{2\sigma_A^2 D^*_1}} \]

as reported by Edwards and Lamkey (2002) and equivalent to the report of Cornelius (1988).

By definition, A and D are independent in noninbred individuals and the correlations between G and A and D were calculated as:

\[ r_{G,A} = \frac{\sigma_A^2}{\sqrt{\sigma_A^2 (\sigma_A^2 + \sigma_D^2)}} \]

and

\[ r_{G,D} = \frac{\sigma_D^2}{\sqrt{\sigma_D^2 (\sigma_A^2 + \sigma_D^2)}} \]

Predicted genetic variances and covariances were calculated as a linear combination of the estimated genetic covariance parameters. The standard errors for the predictions were calculated as a linear function of the asymptotic standard errors of the genetic covariance parameters.
Results

Genetic Variances

All five genetic main effect covariance parameters were significantly different from zero for all traits (Table 3) with the exception of $D_1$ and $H^*$ for mid-pollen and $\sigma^2_D$ and mid-silk. Grain yield differed from all other traits in that the estimate of dominance variance was larger than the estimate of additive variance. The covariance between additive effects and homozygous dominance effects ($D_1$) was negative for all traits with the exception of the flowering traits. Furthermore, $H^*$ was significantly greater than zero for all traits, although the standard errors were relatively large. For grain yield, the genotype by environment (GxE) interaction components were all significantly different from zero and accounted for a large portion of the total genetic variance. The GxE components for the plant height traits and days to mid-silk were relatively small in comparison to the other components.

In all environments with the exception of Rippey, the additive variance was less than the dominance variance for grain yield (Fig. 2). The additive variance for plant height was larger than the dominance variance at all locations with the exception of Carroll 2004 where the two variances were nearly equal. For both traits, the relative values of the parameters showed variation across environments. The Rippey environment appears to be unique in that it varies considerably more from the other five environments as it has a negative value of $H^*$ for plant height and contains extreme values for $D_2^*$ and $D_1$ for grain yield.

Genetic variances were predicted for each generation based on the estimates of the genetic covariance parameters and compared to the observed genetic variances for each generation (Fig. 3). The model accurately predicted variances of both outbred generations for all traits. The model differed most often from the observed values in the S3 and S1
generation. The only difference larger than two standard deviations between the observed and predicted values was for grain moisture in the S3 generation, where the predicted value was 5.04 g kg\(^{-1}\) and the observed value was 3.93 g kg\(^{-1}\).

**Correlations**

For non-inbred individuals, the correlation between G and A was greater than the correlation between G and D for all traits with the exception of grain yield (Table 4). The correlation between G and A for grain yield was 0.65, whereas this correlation ranged from 0.78 to 0.84 for the remaining traits. Grain yield also differed in the correlation between G and D in non-inbred individuals as all other traits were between 0.55 and 0.62, but the correlation for grain yield was 0.75. Correlations between A and D in inbred individuals were found to be negative for grain moisture, grain yield, plant height, and ear height with a range of -0.31 for plant height to -0.59 for ear height. For inbred individuals, silk date was unique as the correlation of G with both A and D was greater than 0.70. In addition, silk date also had a high positive correlation between A and D in inbred individuals.
Discussion

Comparison to BS13

The mating design used in this study was developed based on previous suggestions made by several researchers. Lynch (1988) suggested that a large (over 100) number of lines be evaluated and an effort be made to accurately estimate the additive genetic variance. Cornelius (1988) recommended that outbred progenies from early in the inbreeding process be used to separate the additive and dominance components of genetic variance. We have satisfied Lynch’s (1988) suggestion by evaluating 200 lines in each of five generations, and we have satisfied the second suggestion by testing the S1 topcross generation (S1-TC) as this will give us a clean estimate of the additive genetic variance. Edwards and Lamkey (2002) suggested including outbred progenies of inbred generations as a means of reducing correlations between parameters. Thus we included the S5 topcross generation (S5-TC). In actuality, the inclusion of this generation serves another purpose as well. By inclusion of both the S5 and S5-TC generations we were able to separate inbred dominance deviations from inbred breeding values. To our knowledge, this has not been previously done in any maize population. In all, we evaluated 200 lines in five generations: S1, S3, S5, S1-TC, and S5-TC (Fig. 1).

Edwards and Lamkey (2002) reported genetic covariance estimates for the maize population BS13(S)C0. It was derived from a maize population that was subjected to seven cycles of half-sib selection with Ia13 used as a tester (see Edwards and Lamkey, 2002 for details concerning formation of BS13(S)C0). These estimates were the first significant published estimates of D2* and H* in any crop species. Their breeding design differed from the current design in that they evaluated four inbred generations as well as outbred progeny
from the S1 generation. When compared to estimates obtained from BSCB1, the relative proportions of the genetic covariance parameters are strikingly similar (Table 5). In terms of grain yield, more genetic variance was found in BSCB1, but there was also more inbreeding depression found in BSCB1. What is of more importance is the relative values of $\sigma_A^2$, $\sigma_D^2$, and $D_1$. In both populations, $\sigma_D^2$ was larger than $\sigma_A^2$ and $D_1$ was negative. The absolute value of $D_1$ was nearly half the value of $\sigma_A^2$ in both populations. Thus it appears that for grain yield, the genetic variance structures of these two populations are similar. For grain yield, grain moisture, plant height, and ear height, BSCB1 showed more inbreeding depression as the estimate of $H^*$ was greater. The estimates of $D_2*%$ give an indication of the variance of dominance deviations of inbred lines, i.e. the difference between inbred per se performance and performance of outbred progeny of the same inbred line. BS13 showed relatively more variation in inbred dominance deviations for plant height, ear height, and grain moisture than BSCB1 (Table 5).

In both of these populations, grain yield and grain moisture were the two traits under selection with a greater emphasis placed on grain yield. It appears that selection for grain yield may have resulted in a similar genetic structure in both populations. BS13(S)C0 was initially put through seven cycles of half-sib selection and BSCB1(R)C13 has undergone 13 cycles of reciprocal improvement with BSSS. Given these results for grain yield, it appears that the similar genetic structures may have been produced via selection in two populations that differ in their genetic background. This leads to the speculation that selection for grain yield has resulted in consistent changes in the genetic structure of the two populations. Of greatest concern are the large negative values of $D_1$ found in both populations for grain yield.
A large negative value of $D_1$ indicates that simultaneous improvement of both inbred and outbred performance will be difficult. What would be of great interest in future studies is data on the genetic covariance parameters in an unselected maize population as it is impossible to separate the effects of genetic drift from selection in these two populations.

**Design Issues**

The breeding design used in this study was developed based on suggestions made by several researchers. The S1 generation was included in the design based on suggestions made by Cockerham (1983), because progenies from early generations of inbreeding provide information to estimate $\sigma_D^2$. We also included outbred progenies of the S1 generation to separate the $\sigma_A^2$ and $\sigma_D^2$ components of genetic variance and in essence separate breeding values and dominance deviations in noninbred individuals. Edwards and Lamkey (2002) suggested including outbred progenies of inbred individuals to separate inbred dominance deviations and inbred breeding values. Based on these suggestions we also included S5 progenies and progenies from outbred S5 individuals. The S3 generation was included to provide more information on $H^*$, as it only contributes to the variance of inbred generations and the covariances between inbred generations (Table 1). A shortcoming of our design is that we do not have independent estimates of all five parameters which resulted in multicollinearity among the parameters. We have quantified this multicollinearity via a correlation matrix for grain yield computed from the asymptotic covariance matrix of the estimates:
Our experiment, when compared to the Edwards and Lamkey (2002) experiment, had smaller correlations between parameter estimates in all cases. Although the correlations are reduced from previous designs, our estimator of \( D_1 \) was still correlated with estimators of \( \sigma_A^2 \) and \( D_2^* \). In our design, the coefficients on \( D_1 \) change simultaneously with the coefficients on \( \sigma_A^2 \) and \( D_2^* \) in the same direction:

\[
\begin{align*}
Var(S_1) &= \sigma_A^2 + 1D_1 + 0.125D_2^* + 0.25\sigma_D^2 \\
Var(S_2) &= 1.94\sigma_A^2 + 3.81D_1 + 0.945D_2^* + 0.015H^* + 0.016\sigma_D^2 \\
Var(S_3) &= 2\sigma_A^2 + 4D_1 + 1D_2^*.
\end{align*}
\]

Thus estimators of \( D_1 \) are correlated with estimates of \( \sigma_A^2 \) and \( D_2^* \). New designs should seek to reduce the multicollinearity even further by attempting to either obtain a clean estimate of \( D_1 \) or to break up the dependency between coefficients of \( D_1 \) with both \( \sigma_A^2 \) and \( D_2^* \).

**Role of Dominance**

Previous estimates of \( \sigma_A^2 \) and \( \sigma_D^2 \) in maize have been reported in numerous studies and the general result is that \( \sigma_A^2 \) is larger than \( \sigma_D^2 \) and many times much larger. This has been found to be true for all traits in maize (see Hallauer and Miranda, 1988). The only consistent exception is grain yield in the BSSS population where \( \sigma_D^2 \) is approximately equal to \( \sigma_A^2 \) (Hallauer and Miranda, 1988). It was surprising that we found such a large proportion
of dominance variance in BSCB1 for grain yield as previous estimates have shown additive variance to be larger than dominance variance (Hallauer, 1970). Hallauer (1970) estimated the additive and dominance variances for cycle zero and cycle four of BSCB1. He found the dominance to additive variance ratio to be 0.40 in cycle zero and 0.77 in cycle four, although no statistical test is available for the ratio itself, given the large standard errors associated with the respective variance components the difference between these two ratios is most likely insignificant. We have evaluated cycle 13 of BSCB1 and estimated the dominance to additive variance ratio to be 1.34, a great deal larger than the estimates obtained from cycles zero and four. Although our study and Hallauer's study estimated variance components using different methods, it appears as though dominance variance is increasing in BSCB1 relative to the additive variance.

We can only speculate as to why dominance variance is becoming more prominent in BSCB1, but several explanations do exist. If overdominant gene action is the cause, dominance variance would be larger than additive variance given recessive allele frequencies ranged between 0.2 and 0.8. Given the large amount of evidence against the overdominant hypothesis, we argue that this explanation is not sufficient. Under the dominance hypothesis (Crow, 2000) and a recessive allele frequency of less than 0.35, one would also observe more dominance variance than additive variance. Although this may in fact be the case in BSCB1, it is an unlikely scenario given that recessive alleles with large phenotypic effects are eventually purged via selection. We argue that the most likely reason for the observed variance structure is the accumulation of repulsion phase linkages due to selection.

Commercial maize breeders rarely employ recurrent selection and most often utilize F₂ populations to develop new inbreds. The use of F₂ populations has several advantages
with regards to the genetic covariance parameters. When allele frequencies are equal, $D_1$ and $D_2^*$ both equal zero and when there are only two alleles per locus, $H^*$ equals $\sigma_D^2$ (Cockerham, 1983). When $D_1$ equals zero, the additive and dominance effects within an individual are no longer correlated. The commercial maize industry is primarily concerned with the performance of an individual in a cross to an unrelated individual. In such crosses, $D_1$, $D_2^*$, and $H^*$ only influence inbred per se performance as homozygous dominance deviations are assumed to be absent from a cross of two unrelated individuals. The genetic covariance parameters, however, are a concern to commercial plant breeders as they completely describe the covariance between inbred and outbred performance which is a concern if the inbred is used as a female seed parent.
References


<table>
<thead>
<tr>
<th>Covariance</th>
<th>( \sigma_A^2 )</th>
<th>( \sigma_D^2 )</th>
<th>D1</th>
<th>D2*</th>
<th>H*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Cov}(S_1, S_1) )</td>
<td>1.000</td>
<td>0.250</td>
<td>1.000</td>
<td>0.125</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_1, S_3) )</td>
<td>1.000</td>
<td>0.063</td>
<td>1.370</td>
<td>0.219</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_1, S_2) )</td>
<td>1.000</td>
<td>0.016</td>
<td>1.469</td>
<td>0.242</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_1, S_{1-TC}) )</td>
<td>0.500</td>
<td>0.000</td>
<td>0.250</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_1, S_{5-TC}) )</td>
<td>0.500</td>
<td>0.000</td>
<td>0.250</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_3, S_3) )</td>
<td>1.750</td>
<td>0.063</td>
<td>3.250</td>
<td>0.781</td>
<td>0.047</td>
</tr>
<tr>
<td>( \text{Cov}(S_3, S_5) )</td>
<td>1.750</td>
<td>0.016</td>
<td>3.344</td>
<td>0.805</td>
<td>0.012</td>
</tr>
<tr>
<td>( \text{Cov}(S_3, S_{1-TC}) )</td>
<td>0.500</td>
<td>0.000</td>
<td>0.438</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_3, S_{5-TC}) )</td>
<td>0.875</td>
<td>0.000</td>
<td>0.438</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_5, S_5) )</td>
<td>1.940</td>
<td>0.016</td>
<td>3.810</td>
<td>0.945</td>
<td>0.015</td>
</tr>
<tr>
<td>( \text{Cov}(S_5, S_{1-TC}) )</td>
<td>0.500</td>
<td>0.000</td>
<td>0.484</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_5, S_{5-TC}) )</td>
<td>0.969</td>
<td>0.000</td>
<td>0.484</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_{1-TC}, S_{1-TC}) )</td>
<td>0.250</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_{1-TC}, S_{5-TC}) )</td>
<td>0.250</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \text{Cov}(S_{5-TC}, S_{5-TC}) )</td>
<td>0.484</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Table 2. Expectations \([E(\bullet)]\), variances \([V(\bullet)]\), and covariances \([C(\bullet)]\) for genotypic values \((G)\), breeding values \((A)\), and dominance deviations \((D)\) of noninbred \((F=0)\) and inbred \((F=1)\) individuals from Edwards and Lamkey (2002).

<table>
<thead>
<tr>
<th>Value</th>
<th>Non-inbred</th>
<th>Inbred</th>
</tr>
</thead>
<tbody>
<tr>
<td>(G)</td>
<td>(\alpha_j + \delta_y)</td>
<td>(2\alpha_i + \delta_y)</td>
</tr>
<tr>
<td>(A)</td>
<td>(\alpha_i + \alpha_j)</td>
<td>(2\alpha_i)</td>
</tr>
<tr>
<td>(D)</td>
<td>(\delta_y)</td>
<td>(\delta_i)</td>
</tr>
<tr>
<td>(E(G))</td>
<td>0</td>
<td>(\sum_i p_i \delta_u)</td>
</tr>
<tr>
<td>(E(A))</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(E(D))</td>
<td>0</td>
<td>(\sum_i p_i \delta_u)</td>
</tr>
<tr>
<td>(V(G))</td>
<td>(\sigma_A^2 + \sigma_D^2)</td>
<td>(2\sigma_A^2 + 4D_1 + D_2^*)</td>
</tr>
<tr>
<td>(V(A))</td>
<td>(\sigma_A^2)</td>
<td>(2\sigma_A^2)</td>
</tr>
<tr>
<td>(V(D))</td>
<td>(\sigma_D^2)</td>
<td>(D_2^*)</td>
</tr>
<tr>
<td>(C(G,A))</td>
<td>(\sigma_A^2)</td>
<td>(2\left(\sigma_A^2 + D_1\right))</td>
</tr>
<tr>
<td>(C(G,D))</td>
<td>(\sigma_D^2)</td>
<td>(2D_1 + D_2^*)</td>
</tr>
<tr>
<td>(C(A,D))</td>
<td>0</td>
<td>(2D_1)</td>
</tr>
</tbody>
</table>
Table 3. Genetic covariance parameters for BSCB1(R) C13 for grain yield, grain moisture, plant height, ear height, days to mid-silk, and days to mid-pollen, for a combined analysis over four locations in 2004 and two locations in 2005.

<table>
<thead>
<tr>
<th>Component</th>
<th>Grain yield</th>
<th>Grain moisture</th>
<th>Plant Height</th>
<th>Ear Height</th>
<th>Mid-pollen</th>
<th>Mid-silk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mg ha(^{-2})</td>
<td>g(^{2}) kg(^{-2})</td>
<td>cm(^{2})</td>
<td>cm(^{2})</td>
<td>days(^{2})</td>
<td>days(^{2})</td>
</tr>
<tr>
<td>(\sigma_{A}^{2})</td>
<td>0.61 ± 0.06</td>
<td>2.78 ± 0.21</td>
<td>138 ± 11</td>
<td>119 ± 9</td>
<td>1.41 ± 0.14</td>
<td>1.58 ± 0.15</td>
</tr>
<tr>
<td>(\sigma_{D}^{2})</td>
<td>0.82 ± 0.15</td>
<td>1.55 ± 0.26</td>
<td>86 ± 18</td>
<td>51 ± 11</td>
<td>0.82 ± 0.31</td>
<td>0.67 ± 0.38</td>
</tr>
<tr>
<td>(D_1)</td>
<td>-0.27 ± 0.05</td>
<td>-0.52 ± 0.12</td>
<td>-19 ± 6</td>
<td>-29 ± 5</td>
<td>0.06 ± 0.08</td>
<td>0.36 ± 0.10</td>
</tr>
<tr>
<td>(D_2^*)</td>
<td>0.87 ± 0.14</td>
<td>1.36 ± 0.22</td>
<td>53 ± 10</td>
<td>40 ± 8</td>
<td>0.64 ± 0.18</td>
<td>1.14 ± 0.27</td>
</tr>
<tr>
<td>(H^*)</td>
<td>6.21 ± 1.46</td>
<td>14.91 ± 4.51</td>
<td>729 ± 201</td>
<td>407 ± 129</td>
<td>2.25 ± 2.42</td>
<td>12.41 ± 5.12</td>
</tr>
<tr>
<td>(\sigma_{AE}^{2})</td>
<td>0.25 ± 0.02</td>
<td>0.35 ± 0.03</td>
<td>7 ± 2</td>
<td>8 ± 2</td>
<td>0.14 ± 0.04</td>
<td>0.19 ± 0.05</td>
</tr>
<tr>
<td>(\sigma_{DE}^{2})</td>
<td>0.22 ± 0.07</td>
<td>0.58 ± 0.11</td>
<td>-1 ± 9</td>
<td>7 ± 7</td>
<td>-0.29 ± 0.17</td>
<td>0.00 ± 0.24</td>
</tr>
<tr>
<td>(D_{1E})</td>
<td>-0.16 ± 0.03</td>
<td>-0.08 ± 0.03</td>
<td>-2 ± 2</td>
<td>-4 ± 2</td>
<td>-0.13 ± 0.05</td>
<td>-0.06 ± 0.06</td>
</tr>
<tr>
<td>(D_{2E}^*)</td>
<td>0.40 ± 0.07</td>
<td>0.39 ± 0.09</td>
<td>5 ± 6</td>
<td>11 ± 5</td>
<td>0.47 ± 0.16</td>
<td>0.41 ± 0.20</td>
</tr>
<tr>
<td>(H_{E}^*)</td>
<td>2.46 ± 0.72</td>
<td>21.5 ± 2.01</td>
<td>140 ± 82</td>
<td>60 ± 52</td>
<td>1.13 ± 1.66</td>
<td>0.38 ± 2.87</td>
</tr>
</tbody>
</table>
Table 4. Correlations between genotypic values, breeding values, and dominance deviations in both inbred and non-inbred individuals in BSCB1(R)C13.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Non-inbred individuals</th>
<th></th>
<th>Inbred individuals</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r(G,A)*</td>
<td>r(G,D)</td>
<td>r(G,A)</td>
<td>r(G,D)</td>
</tr>
<tr>
<td>Grain yield</td>
<td>0.65</td>
<td>0.75</td>
<td>0.61</td>
<td>0.35</td>
</tr>
<tr>
<td>Grain moisture</td>
<td>0.80</td>
<td>0.60</td>
<td>0.87</td>
<td>0.12</td>
</tr>
<tr>
<td>Plant height</td>
<td>0.78</td>
<td>0.62</td>
<td>0.90</td>
<td>0.13</td>
</tr>
<tr>
<td>Ear height</td>
<td>0.84</td>
<td>0.55</td>
<td>0.92</td>
<td>-0.22</td>
</tr>
<tr>
<td>Mid-pollen</td>
<td>0.80</td>
<td>0.61</td>
<td>0.91</td>
<td>0.49</td>
</tr>
<tr>
<td>Mid-silk</td>
<td>0.84</td>
<td>0.55</td>
<td>0.91</td>
<td>0.72</td>
</tr>
</tbody>
</table>

* G= genotypic value, A= breeding value, and D= dominance deviation.
Table 5. Genetic covariance parameters for BSCB1(R)C13 and BS13(S)C0 for grain yield, grain moisture, plant height, ear height, days to mid-silk, and days to mid-pollen, for a combined analysis over six locations.

<table>
<thead>
<tr>
<th>Component</th>
<th>BSCB1(R)C13</th>
<th>BS13(S)C0</th>
<th>BSCB1(R)C13</th>
<th>BS13(S)C0</th>
<th>BSCB1(R)C13</th>
<th>BS13(S)C0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain yield</td>
<td>Mg m⁻² ha⁻²</td>
<td>Grain moisture</td>
<td>g⁻³ kg⁻²</td>
<td>Plant Height</td>
<td>cm²</td>
<td>Ear Height</td>
</tr>
<tr>
<td>σ²_A</td>
<td>0.61 ± 0.06</td>
<td>0.29 ± 0.05</td>
<td>138 ± 11</td>
<td>208 ± 23</td>
<td>1.6 ± 0.2</td>
<td>4.1 ± 0.5</td>
</tr>
<tr>
<td>σ²_D</td>
<td>0.82 ± 0.15</td>
<td>0.32 ± 0.09</td>
<td>86 ± 18</td>
<td>64 ± 21</td>
<td>0.7 ± 0.4</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>D1</td>
<td>-0.27 ± 0.05</td>
<td>-0.18 ± 0.06</td>
<td>-19 ± 6</td>
<td>-76 ± 18</td>
<td>0.4 ± 0.1</td>
<td>-1.0 ± 0.4</td>
</tr>
<tr>
<td>D2*</td>
<td>0.87 ± 0.14</td>
<td>0.85 ± 0.19</td>
<td>53 ± 10</td>
<td>194 ± 47</td>
<td>1.1 ± 0.3</td>
<td>1.7 ± 1.0</td>
</tr>
<tr>
<td>H*</td>
<td>6.21 ± 1.46</td>
<td>1.55 ± 0.48</td>
<td>729 ± 201</td>
<td>661 ± 149</td>
<td>12.4 ± 5.1</td>
<td>21.0 ± 4.2</td>
</tr>
<tr>
<td>Grain moisture</td>
<td>g⁻³ kg⁻²</td>
<td>Plant Height</td>
<td>cm²</td>
<td>Ear Height</td>
<td>cm³</td>
<td>Mid-silk</td>
</tr>
<tr>
<td>σ²_A</td>
<td>2.75 ± 0.21</td>
<td>5.20 ± 0.60</td>
<td>119 ± 9</td>
<td>149 ± 17</td>
<td>1.4 ± 0.1</td>
<td>2.1 ± 0.3</td>
</tr>
<tr>
<td>σ²_D</td>
<td>1.55 ± 0.26</td>
<td>1.70 ± 0.70</td>
<td>51 ± 11</td>
<td>44 ± 14</td>
<td>0.8 ± 0.3</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>D1</td>
<td>-0.52 ± 0.12</td>
<td>-0.40 ± 0.40</td>
<td>-29 ± 5</td>
<td>-66 ± 13</td>
<td>0.1 ± 0.1</td>
<td>-0.3 ± 0.3</td>
</tr>
<tr>
<td>D2*</td>
<td>1.36 ± 0.22</td>
<td>2.90 ± 1.20</td>
<td>40 ± 8</td>
<td>147 ± 33</td>
<td>0.6 ± 0.2</td>
<td>0.6 ± 0.7</td>
</tr>
<tr>
<td>H*</td>
<td>14.91 ± 4.51</td>
<td>6.60 ± 4.80</td>
<td>407 ± 99</td>
<td>344 ± 99</td>
<td>2.3 ± 2.4</td>
<td>6.4 ± 2.1</td>
</tr>
</tbody>
</table>

† This standard error was incorrectly reported in Edwards and Lamkey (2002), 0.19 is the correct standard error.
Figure Captions

Figure 1. A diagram representing the mating design used for estimation of the genetic covariance parameters. Ovals encompass the generations used in the analysis.

Figure 2. Values of the $\sigma^2_D$, $D_1$, and $D_2^*$ expressed as a relative percentage of the respective $\sigma^2_A$ for six environments and the combined analysis for grain yield (A) and plant height (B).

Figure 3. Line graphs representing the comparison of observed genetic variances and predicted genetic variances using the estimated genetic covariance parameters for plant height, ear height, grain yield, grain moisture, days to mid pollen (pollen date) and days to mid silk (silk date). Standard error bars represent one standard deviation.
Figure 1

BSCB1(R)C13

\[ \begin{align*}
& \rightarrow 200 S_1 \\
& \rightarrow 200 S_2 \\
& \rightarrow 200 S_3 \\
& \rightarrow 200 S_4 \\
& \rightarrow 200 S_5 \\
& \end{align*} \]

BSCB1(R)C13

\[ \begin{align*}
& 200 S_1 - TC \\
& 200 S_5 - TC \\
& \end{align*} \]
Figure 2

A) Dominance Variance, D1, D2*

B) Dominance Variance, D1, D2*
Chapter 4. The Relationship Between Genetic Variance Components and Predicted Gains From Selection

A paper to be submitted to Crop Science.

Brandon M. Wardyn, Jode W. Edwards* and Kendall R. Lamkey

Abstract

Heritability is a core component to the genetic gain equation and is one of, if not the most, frequently estimated and discussed quantitative genetic parameters. We have defined the heritability as the covariance between selection units and response units under a clearly defined reference population of both genotypes and environments. Many previous estimates of genetic gain have ignored the effects of both inbreeding and dominance on heritability. We have used a model that accounts for these effects and have obtained a heritability estimate that contains less bias than previous estimates. Genetic gain equations were calculated for three selection units each associated with four response units. Our results indicated that the bias introduced by the combination of inbreeding and dominance inflated genetic gain predictions for some traits. In addition, the bias was erroneously suggesting that inbred progeny selection methods were superior to outbred progeny selection methods. We found that for traits with a significant amount of dominant gene action, outbred progeny selection was superior to inbred progeny selection in the presence of inbreeding.
**Introduction**

The common goal of plant breeding programs revolves around shifting a mean in a desired direction via exploitation of the genetic differences between individuals due to the laws of segregation and independent assortment. Without phenotypic variance among individuals, the skills of a plant breeder would be ineffective. The genetic component of the phenotypic variance is of primary interest to plant breeders due to its repeatability over an infinite set of environments. This genetic component of a population, or stated differently the genetic structure of a population, is determined by allele frequencies and gene effects. Although genetic variance among individuals is a necessity for progress in selection programs, it is not exclusive nor is it the sole indicator of the rate of change that can be made from selection.

For selection to be effective, a covariance between relatives due to genetic effects must exist. The success of selection among individuals thus depends on the heritability (Lush, 1936). Hanson (1963) defined heritability as “the fraction of the selection differential expected to be gained when selection is practiced on a defined reference unit”. In a selection framework, the genetic covariance between relatives is the genetic covariance between the selection unit and the response unit. It is intuitive, that any selection scheme involving plants must have both a clearly defined selection unit as well as a clearly defined response unit. Furthermore, to adequately define the context of a heritability estimate, a randomly sampled reference population of genotypes and reference population of environments must be defined (Dudley and Moll, 1969; Nyquist 1991, pp. 239-243). The concept of heritability is not new and has been reviewed and estimated extensively in the literature (Nyquist, 1991), and it is
most commonly defined as the “the extent to which phenotypes are determined by genes transmitted from the parents” (Falconer and Mackay pp. 123, 1996).

If a constant amount of genetic variance among a reference population of genotypes is assumed, the only ways to increase heritability are to increase the accuracy and precision of the estimators of genotypic values or to maximize the covariance between the selection unit and the response unit. Obtaining precise and accurate estimates of genotypic values is a fundamental part of any scientific experiment and merits no further discussion. Thus, only the manipulation of the covariance between the selection unit and response unit is at the discretion of the breeder. Historically, breeders have manipulated this covariance via selection methods. Fehr (1987) outlines several classical selection methods that have historically been a popular choice among plant breeders. In addition, numerous variations and modifications have been introduced by individuals to fit their budget, time frame, or biology of the target species. The choice of selection method most often is determined by the amount of gain that is expected per year based on quantitative genetic theory. What has commonly been observed, however, is that the realized gain often falls short of predicted gain (Weyhrich et al., 1998; Holland et al., 2003).

One hypothesis for the discrepancy between predicted and observed rates of gain is that there is a bias involved in the standard theory used to predict genetic gain. The bias is introduced when the family variance or additive variance is used as the numerator in the heritability equation. Heritability estimated in this way is not equal to the covariance between the selection unit and response unit (Holland et al., 2003; Lamkey and Hallauer, 1987). Much of the literature on selection theory has either ignored this bias or deemed it negligible. Hallauer and Miranda (1988) published a general gain equation which included a
bias term that was defined as "the deviation from the additive genetic variance". Since estimates of this bias were often unestimable without special experimental designs, the additive variance was often used as the numerator of the heritability equation. Heritability defined in this manner was, and often still is, referred to as the narrow-sense heritability (Falconer and Mackay, 1996). This, however, is an oversimplification of heritability. By using only the additive variance in the heritability estimate, it is consequently assumed that dominance variance is absent. This assumption, primarily in maize for grain yield, has been shown to be invalid as dominance variance has been estimated to be larger than the additive variance (Edwards and Lamkey, 2002; Wardyn et al., 2006).

Standard theory has for the most part ignored dominance variance with inbred progeny selection methods (Falconer and Mackay, 1996; Mather and Jinks, 1977). Comstock (1964) recommends inbred progeny selection over all other selection methods, as he estimated inbred progeny selection to be twice as effective as half-sib selection. Horner (1969) also recommends inbred progeny selection based on the relative proportion of additive variance displayed among the progeny means. Inbred progeny selection, under a completely additive scenario will always be superior to other forms of selection. In the presence of dominant gene action, however, inbred progeny selection becomes less advantageous to outbred progeny selection methods. Inbred progeny selection has also failed to improve the Iowa Stiff Stalk synthetic population after six cycles of S2 selection for grain yield, which is a drastic departure from standard theory given the level of additive dominance in Iowa Stiff Stalk Synthetic (Lamkey, 1992).

These discrepancies between predicted and realized gains emphasize the importance of accurately defining the heritability as the covariance between the selection unit and the
response unit without restrictions on the types of gene action. By defining the heritability in this way, predicted gains from selection will be unbiased. We have set up an experiment in which we will be able to answer the following questions: i) How do the predicted gain equations change when the covariance between response units and selection units is used to calculate heritability? ii) Which of the genetic covariance parameters influence genetic gain? iii) To what degree do predicted gains change when empirical estimates of the genetic covariance parameters are used to predict gain?
Materials and Methods

Germplasm & Mating Design

The germplasm source for this study consisted of two maize populations developed at Iowa State University: BSCB1RC(13), a member of the non-stiff stalk heterotic pattern, and BS13SC(0), a member of the stiff stalk heterotic pattern. BSCB1RC(13) had undergone 13 cycles of reciprocal recurrent selection and BS13SC(0) had undergone seven cycles of half-sib selection with Ia13 as a tester. From each population, 200 noninbred individuals were sampled and both inbred and outbred progeny were obtained from each individual. Refer to Wardyn et al. (2006) for details concerning progeny formation in BSCB1RC(13) and Edwards and Lamkey (2002) for details concerning BS13SC(0).

Genetic Model

The same genetic model used by Wardyn et al. (2006) and Edwards and Lamkey (2002) was used for this study. The model is based on an extension of Fisher’s (1918) genetic model by Harris (1964) to include inbred relatives. Harris developed a completely general parameterization of the covariance between two individuals of a population with an arbitrary relationship with each other and with arbitrary levels of inbreeding. The beauty of Harris’s model is that it completely describes the covariance between two individuals and thus it can be used to model either a linear or non-linear relationship between inbreeding and genetic variance. It should be noted that this general parameterization is true under the following assumptions i) there is no linkage among the loci that influence the traits being evaluated, ii) the original population is random mating, iii) there has been no selection practiced during the development of the two individuals, and iv) the individuals have
autosomal diploid loci. Using this model and Cockerham's (1983) notation, the covariance between two individuals (X and Y) can be represented by:

$$Cov(X, Y) = 2\theta_{XY}\sigma_A^2 + 2(\Delta_{XY} - \delta_{XY})\sigma_D^2 + 2(\gamma_{XY} + \gamma_{X\bar{Y}})D_1 + \delta_{XY}D_2^* + \left(\Delta_{\bar{X},\bar{Y}} - F_{\bar{X}}F_{\bar{Y}}\right)H^*$$

where:

- $\sigma_A^2$ = additive genetic variance
- $\sigma_D^2$ = dominance variance
- $D_1$ = covariance between additive effects and homozygous dominance effects
- $D_2^*$ = the variance of homozygous dominance deviations
- $H^*$ = the sum of homozygous dominance deviations, squared,

which are collectively referred to as genetic covariance parameters. Cockerham (1971) describes the calculation of the 15 probabilities of identity by decent for two, three, and four alleles which are jointly represented by $\theta_{XY}, \Delta_{XY}, \delta_{XY}, \gamma_{XY}, \gamma_{X\bar{Y}}, \Delta_{\bar{X},\bar{Y}}, F_{\bar{X}},$ and $F_{\bar{Y}}$. The 15 probabilities determine the coefficients for each of the five genetic covariance parameters ($\sigma_A^2, \sigma_D^2, D_1, D_2^*, H^*$). The model can be extended to include genotype by environment effects and higher forms of epistasis as well. Epistasis was assumed to be negligible in BS13SC(0) (Edwards and Lamkey, 2002) and additive by additive epistasis was found to be insignificant in BSCB1RC(13) (unpublished data). Thus we proceeded under the assumption that any additional forms of epistasis in BSCB1RC(13) would be insignificant as well.

**Estimation**
Estimates of the genetic covariance parameters and their associated standard errors were obtained via a mixed linear model. Genotype by environment interactions were accounted for in the analysis as each population was evaluated in six environments each with two replications. Refer to Wardyn et al (2006) and Edwards and Lamkey (2002) for details concerning the estimation procedure and experimental design for BSCB1RC(13) and BS13SC(0), respectively. Estimates of the genetic covariance parameters were obtained for grain yield (Mg ha\(^{-1}\); adjusted to 15.5% moisture) and plant height (cm). Plant height was measured as the distance from the soil surface to the uppermost leaf collar.

The resulting genetic covariance estimates were inserted into genetic gain equations to predict the genetic gain with differing selection and response units. Genetic gain equations were calculated using the procedure described in Empig et al. (1981). This method is based on the following assumptions: the organism is diploid, no linkage, two alleles per locus and no epistasis. The formation of the genetic gain equations can be carried out in three steps: (1) calculate the change in gene frequency at a locus in the improved population that was formed from the recombination of selected individuals; (2) calculate the change of the mean of the improved population; (3) multiply (1) and (2). See Empig et al. (1981) for further details. As a means of verification, the gain equations were also calculated using the higher-order measures of identity by descent developed by Cockerham (1971; 1983). The verification process directly calculated the covariance between the response unit and the selection unit for the equation:

\[
\Delta G = \frac{s}{\sigma_p^2} \left[ \text{Cov}(\text{Response Unit}, \text{Selection Unit}) \right],
\]

where \(s\) = selection differential and \(\sigma_p^2\) = the phenotypic variance of the response unit.
**Results and Discussion**

**Model Assumptions**

The assumptions associated with Harris’s (1964) model are: i) there is no linkage among the loci that influence the traits being evaluated, ii) the original population is random mating, iii) there has been no selection practiced during the development of the two individuals, and iv) the individuals have autosomal diploid loci. Both BS13SC(0) and BSCB1RC(13) are random mating populations with autosomal diploid loci. Individuals were randomly selected from the base populations and no selection was applied during the formation of the inbred progenies. Only the assumption of linkage equilibrium is of concern in these two populations. After the recombination of selected individuals, each population was random mated for two generations. While we do not argue that this small amount of random mating was sufficient to break up all linkages, we proceeded under the assumption of no linkage. The concept of linkage has, for the most part, been ignored in the standard theory previously used to predict gain. It is our intention to use Harris’s (1964) model with the assumption of no linkage, but with the caveat being that linkage is most likely present but to an unknown degree. This is a form of bias in our model which resembles the bias in early quantitative genetic theory which ignored the effects of dominance.

**Heritability**

A common approach to estimating heritability in the literature is to regress the phenotypic values of offspring on parental values. Under a completely additive model, this approach is correct and yields the true heritability. When the genetic model deviates from complete additivity, the parent-offspring regression is most often biased upward. The cause for this bias is most often the genetic covariance parameters introduced by the combination
of inbreeding and dominance. Cockerham (1983) outlines the model that accurately estimates the covariance between two individuals regardless of the level of dominance or inbreeding level. Using Cockerham’s (1983) notation, inbreeding and dominance introduce three new parameters to the genetic model that have historically not been accounted for:

- **D1** the covariance between additive effects and inbred dominance effects,
- **D2** the variance of homozygous dominance deviations,
- **H** the square of inbreeding depression.

Holland et al. (2003) discusses some of the implications of these three parameters on heritability estimation. The effects of D1, D2*, and H* on heritability are directly related to the amount of inbreeding and dominance in a population. If there is no dominant gene action, D1, D2*, and H* are all equivalent to zero. Thus, dominance is needed for the contributions of D1, D2*, and H* to the genetic model. The effect of inbreeding can be seen by comparing the coefficients of the genetic covariance parameters used to calculate the variances among individuals (Table 6). As individuals become more inbred, the contributions of the D1 and D2* account for a larger portion of the total genetic variance within a generation. D1 and D2* also contribute to the covariances among inbred generations as well (Table 6). As in the case of generation variances, the effects of D1, D2*, and H* become more prevalent as the individuals become more inbred (Table 6).

Heritability estimated from these covariances will be unbiased in the absence of epistasis and linkage. Epistasis can be fit in this model and an analysis was performed fitting the additive by additive variance in BSCB1RC(13) (unpublished data). The additive by additive variance was found to be insignificant in BSCB1RC(13), however a high degree of multicollinearity among the parameters was also found. Future studies should seek breeding
designs that avoid the multicollinearity among parameters that developed by including epistasis in this breeding design. If epistasis is truly absent from the model, the covariances among generations will be an unbiased estimate of the heritability in the absence of linkage. The genetic covariance parameters also influence the genetic variance among individuals which have undergone some form of inbreeding, or stated differently, the denominator of the heritability equation. In a properly executed experiment, the genetic variance should constitute a large proportion of the phenotypic variance. It follows that any effects of the genetic covariance parameters on genetic variance will be mirrored in the estimate of phenotypic variance. Thus in the presence of inbreeding and dominance, the genetic covariance parameters influence heritability through the phenotypic variance of the response units (denominator of the heritability equation) and the covariance between the response units and selection units (numerator of the heritability equation).

**Impacts on Selection**

The foundations of any breeding program are dependent upon the units in which selection acts and the units in which response is measured. Genetic gain is defined as the gain in performance of the response unit, which most often in maize is outbred performance. The response from different selection units depends on first, the amount of observed genetic variance among the selection units, and second, the genetic variance structure of the observed genetic variance. The covariance between the selection unit and the response unit is a linear function of the genetic covariance parameters (Holland, 2003). This covariance is a component of the genetic gain equation which follows the general form:

\[ \Delta G = \frac{s}{\sigma_p^2} \text{Cov(Response Unit, Selection Unit)} \]
where $\Delta G$ is the genetic gain per cycle of selection, $s$ is the selection differential, and $\sigma_r^2$ is the phenotypic variance of the selection unit.

Three common selection units used in maize are S1, S2 and half-sib progenies (the common parent of each half-sib progeny is the actual selection unit for half-sib selection). Each selection unit contains a different genetic gain formula with respect to a specified response unit (Table 7). The gain equations presented in Table 7 represent the expected genetic gain using either S1, S2, or half-sib families as the selection unit and one of four response units. We define the immediate response as the gain in performance realized by re-evaluating the selected entries, the S0 response is the response in the improved outbred population (here the improved population is defined as the population formed by recombining the selected individuals or families), the S1 and S2 responses are the improvement in the S1 and S2 lines developed from the improved population, respectively. In these equations we assume: remnant seed of the selected S1 lines are recombined in S1 selection, remnant seed of the selected S2 lines are recombined in S2 selection, and remnant seed of the common parent with an inbreeding coefficient of $F$ is recombined in half-sib selection. From the gain equations, it is clear that any selection method involving inbred progeny in either the response or selection unit will be influenced by parameters other than $\sigma_r^2$. The degree to which $D_1$ and $D_2^*$ influence the gain from selection is a function of the covariance between the selection unit and the response unit. Perhaps the most important point to be taken from the gain equations is the lack of $D_1$ for half-sib selection when the S0 population is the response unit.
By directly applying Harris's (1964) model, the estimated genetic covariance parameters (Table 8) and their influence on genetic gain can be estimated. The expected gain under three different selection methods for four different response units was estimated for BSCB1RC(13) and BS13SC(0) for both grain yield and plant height (Table 9). Gains under the full model have been calculated using the equations in Table 7 and estimates of the genetic covariance parameters in Table 8. The reduced model utilized the same equations except we used only the additive variance and all other parameters were assumed to equal zero. Predicted gains under the repeatability column used the general equation:

$$AG = s \frac{\sigma_{family}^2}{\sigma_{phenotypic}^2},$$

where the family variance is the genetic variance among the respective families for each selection method. Traditionally, the family and phenotypic variances are the only variance estimates accessible by the plant breeder and thus are used to predict gain.

Using repeatability in the gain equation results in inbred progeny selection being superior to half-sib selection. This result follows standard theory in that as an individual becomes more inbred its associated performance becomes more stable i.e. a higher degree of repeatability. If the full model is used to predict gain, however, half-sib selection is superior whenever the response unit is the improved population. The advantage of half-sib selection is greater for grain yield than plant height in both populations. Plant height contains a larger proportion of additive variance than grain yield in both populations. It follows that plant height would fall in line with the standard theory that assumes a completely additive model. Since grain yield contains a larger proportion of dominance, the bias incurred when dominance is ignored is larger than the bias associated with the same assumptions for plant
height. The effect of this bias is highlighted by the differing results found when the reduced model is used to predict gain. Under the reduced model, S2 selection is superior to either S1 or half-sib selection in both populations for both plant height and grain yield. In most recurrent selection programs, the S0 population is the response unit of interest and improvement of grain yield is the primary goal. Under these circumstances, gain would be severely over predicted in both populations if the reduced model was used and S2 selection was subsequently carried out. This demonstrates the errors that can occur when selection programs are based upon gain equations that do not account for the underlying genetic structure of a population or the true heritability.

Horner et al. (1969) states that in theory, inbred progeny selection would be superior to half-sib selection because the variance among inbred families is larger than the variance among half-sib families due to the higher proportion of additive genetic variance found among inbred families, which is in agreement with Comstock’s (1964) suggestions. While this reasoning is sound, it relies on the assumption that the heritability is identical for all selection methods. This is equivalent to using the family variance as the numerator of the heritability equation in all applications. The gain equations presented in Table 7 use the covariance between the response unit and the selection unit is the numerator of the heritability equation. A comparison of the gain equations reveals that the numerator of the heritability is not equivalent across selection methods and also changes in accordance with the response. Previous work on this issue demonstrated that heritability is dependent upon the response unit, the selection unit and can be empirically estimated with the genetic covariance parameters (Holland et al., 2003). Our empirical estimates of the genetic
covariance parameters allow us to predict selection gains with the correct heritabilities in BSCB1RC(13) and BS13SC(0) for any selection unit and any response unit.

These results provide explanations as to the failure of inbred progeny selection in some applications. Most notably, it provides a partial explanation of the failure of S2-progeny recurrent selection in BS13SC(0) for grain yield (Lamkey, 1992). It does not fully explain the lack of progress in BS13SC(0) with S2 selection as even under the full model, positive gain is predicted in the improved population. These results do, however, give some explanation as to why realized gains usually fall short of predicted gains. The presence of dominance negatively impacts realized gain via a negative D1. The tendency of D1 to be negative implies that gains being made via additive effects are being offset via losses through dominance effects since D1 represents the covariance between additive effects and homozygous dominance effects.

What remains to be answered, however, is if selection has molded the genetic variance structure. All of the estimates in maize of the genetic covariance parameters have shown similar results across populations (Coors, 1988; Edwards and Lamkey, 2002; Wardyn et al., 2006). While the BSCB1RC(13) and BS13SC(0) populations differ in their genetic background, they both have been influenced by selection for several cycles. It is also important to note that BSCB1RC(13) was improved via reciprocal recurrent selection with the population BSSS while BS13SC(0) was improved via half-sib selection with an inbred tester (Ia13). We hypothesize that selection may be accumulating repulsion phase linkages in each population and thus resulting in a large amount of dominance variance and a large and negative D1 estimate. It would be of great interest to obtain estimates of the genetic covariance parameters in a population that was shown to be in linkage equilibrium. If
selection is molding the genetic variance structure via linkage, we find it surprising that relatively the same genetic variance structure has been produced in two populations that differ in both their genetic background and their improvement methods.
References


2nd ed. Iowa State Univ. Press, Ames, IA.


Table 6. Coefficients for genotypic covariance components for the variances of the $S_0$, $S_1$, $S_3$, and $S_5$, generations and the covariance of the $S_1$ generation with the $S_0$, $S_3$, and $S_5$ generations respectively.

<table>
<thead>
<tr>
<th>Covariance</th>
<th>Component</th>
<th>$\sigma^2_A$</th>
<th>$\sigma^2_D$</th>
<th>D1</th>
<th>D2*</th>
<th>H*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var ($S_0$)</td>
<td></td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Var ($S_1$)</td>
<td></td>
<td>1.00</td>
<td>0.25</td>
<td>1.00</td>
<td>0.13</td>
<td>0.00</td>
</tr>
<tr>
<td>Var ($S_3$)</td>
<td></td>
<td>1.75</td>
<td>0.06</td>
<td>3.25</td>
<td>0.78</td>
<td>0.05</td>
</tr>
<tr>
<td>Var ($S_5$)</td>
<td></td>
<td>1.94</td>
<td>0.02</td>
<td>3.81</td>
<td>0.95</td>
<td>0.02</td>
</tr>
<tr>
<td>Cov($S_1$, $S_0$)</td>
<td></td>
<td>0.50</td>
<td>0.00</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Cov($S_1$, $S_3$)</td>
<td></td>
<td>1.00</td>
<td>0.06</td>
<td>1.37</td>
<td>0.22</td>
<td>0.00</td>
</tr>
<tr>
<td>Cov($S_1$, $S_5$)</td>
<td></td>
<td>1.00</td>
<td>0.02</td>
<td>1.47</td>
<td>0.24</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 7. Genetic gain equations for $S_1$, $S_2$, and half-sib selection when the response is measured as the immediate response, the response in the improved population, the response in the improved population of $S_1$ lines, and the response in the improved population of $S_2$ lines.

<table>
<thead>
<tr>
<th>Selection Unit</th>
<th>Immediate</th>
<th>$S_0$</th>
<th>$S_1$</th>
<th>$S_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>$\frac{s}{\sigma_{S1}} \left( \sigma_i^2 + \frac{1}{4} \sigma_i^2 + D_i + \frac{1}{8} D_i^* \right)$</td>
<td>$\frac{s}{\sigma_{S1}} \left( \sigma_i^2 + \frac{1}{2} D_i \right)$</td>
<td>$\frac{s}{\sigma_{S1}} \left( \sigma_i^2 + D_i + \frac{1}{8} D_i^* \right)$</td>
<td>$\frac{s}{\sigma_{S1}} \left( \sigma_i^2 + \frac{5}{4} D_i + \frac{3}{16} D_i^* \right)$</td>
</tr>
<tr>
<td>$S_2$</td>
<td>$\frac{s}{\sigma_{S2}} \left( \frac{3}{2} \sigma_i^2 + \frac{1}{8} \sigma_i^2 + \frac{5}{2} D_i + \frac{9}{16} D_i^* + \frac{1}{16} H^* \right)$</td>
<td>$\frac{s}{\sigma_{S2}} \left( \frac{3}{2} \sigma_i^2 + \frac{5}{4} D_i \right)$</td>
<td>$\frac{s}{\sigma_{S2}} \left( \frac{3}{2} \sigma_i^2 + 2D_i + \frac{5}{16} D_i^* \right)$</td>
<td>$\frac{s}{\sigma_{S2}} \left( \frac{3}{2} \sigma_i^2 + \frac{19}{8} D_i + \frac{15}{22} D_i^* \right)$</td>
</tr>
<tr>
<td>Half-sib</td>
<td>$\frac{s}{\sigma_{Hs}} \left( \frac{1}{4} \sigma_i^2 \right)$</td>
<td>$\frac{s}{\sigma_{Hs}} \left( \frac{(1+F)}{2} \sigma_i^2 \right)$</td>
<td>$\frac{s}{\sigma_{Hs}} \left( \frac{(1+F)}{2} \sigma_i^2 + \frac{(1+F)}{4} D_i \right)$</td>
<td>$\frac{s}{\sigma_{Hs}} \left( \frac{(1+F)}{2} \sigma_i^2 + \frac{3(1+F)}{8} D_i \right)$</td>
</tr>
</tbody>
</table>
Table 8. Genetic covariance parameters for BSCB1(R)C13 and BS13(S)C0 for grain yield and plant height for a combined analysis over six locations.

<table>
<thead>
<tr>
<th>Component</th>
<th>Grain yield Mg² ha⁻²</th>
<th>Plant Height cm²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BSCB1(R)C13</td>
<td>BS13(S)C0</td>
</tr>
<tr>
<td>σ²ₓ</td>
<td>0.61 ± 0.06</td>
<td>0.29 ± 0.05</td>
</tr>
<tr>
<td>σ²₀</td>
<td>0.82 ± 0.15</td>
<td>0.32 ± 0.09</td>
</tr>
<tr>
<td>D¹</td>
<td>-0.27 ± 0.05</td>
<td>-0.18 ± 0.06</td>
</tr>
<tr>
<td>D²*</td>
<td>0.87 ± 0.14</td>
<td>0.85 ± 0.19†</td>
</tr>
<tr>
<td>H*</td>
<td>6.21 ± 1.46</td>
<td>1.55 ± 0.48</td>
</tr>
</tbody>
</table>

† This standard error was incorrectly reported in Edwards and Lamkey (2002), 0.19 is the correct standard error.
Table 9. Predicted genetic gain for $S_1$, $S_2$ and half-sib selection methods based on current estimates of the genetic covariance parameters in BSCB1RC(13) and BS13SC(0) for grain yield and plant height for three methods of calculating heritability: the full genetic model, the reduced genetic model, and repeatability.

### BSCB1RC(13)

<table>
<thead>
<tr>
<th>Selection Unit</th>
<th>Response Unit</th>
<th>Grain yield ($\text{Mg ha}^{-2}$)</th>
<th>Plant Height ($\text{cm}^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Full model</strong></td>
<td><strong>Reduced model</strong></td>
</tr>
<tr>
<td>S1 Immediate</td>
<td></td>
<td>1.42</td>
<td>1.32</td>
</tr>
<tr>
<td>S2 Immediate</td>
<td></td>
<td>1.94</td>
<td>1.45</td>
</tr>
<tr>
<td>HS Immediate</td>
<td></td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>S1 S0 Pop</td>
<td></td>
<td>1.03</td>
<td>1.32</td>
</tr>
<tr>
<td>S2 S0 Pop</td>
<td></td>
<td>0.92</td>
<td>1.45</td>
</tr>
<tr>
<td>HS S0 Pop</td>
<td></td>
<td>1.37</td>
<td>1.37</td>
</tr>
<tr>
<td>S1 S1 Pop</td>
<td></td>
<td>0.97</td>
<td>1.32</td>
</tr>
<tr>
<td>S2 S1 Pop</td>
<td></td>
<td>1.03</td>
<td>1.45</td>
</tr>
<tr>
<td>HS S1 Pop</td>
<td></td>
<td>1.07</td>
<td>1.37</td>
</tr>
<tr>
<td>S1 S2 Pop</td>
<td></td>
<td>0.95</td>
<td>1.32</td>
</tr>
<tr>
<td>S2 S2 Pop</td>
<td></td>
<td>1.08</td>
<td>1.45</td>
</tr>
<tr>
<td>HS S2 Pop</td>
<td></td>
<td>0.96</td>
<td>1.37</td>
</tr>
</tbody>
</table>

### BS13SC(0)

<table>
<thead>
<tr>
<th>Selection Unit</th>
<th>Response Unit</th>
<th>Grain yield ($\text{Mg ha}^{-2}$)</th>
<th>Plant Height ($\text{cm}^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Full model</strong></td>
<td><strong>Reduced model</strong></td>
</tr>
<tr>
<td>S1 Immediate</td>
<td></td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td>S2 Immediate</td>
<td></td>
<td>1.36</td>
<td>0.98</td>
</tr>
<tr>
<td>HS Immediate</td>
<td></td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>S1 S0 Pop</td>
<td></td>
<td>0.64</td>
<td>0.94</td>
</tr>
<tr>
<td>S2 S0 Pop</td>
<td></td>
<td>0.48</td>
<td>0.98</td>
</tr>
<tr>
<td>HS S0 Pop</td>
<td></td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>S1 S1 Pop</td>
<td></td>
<td>0.70</td>
<td>0.94</td>
</tr>
<tr>
<td>S2 S1 Pop</td>
<td></td>
<td>0.77</td>
<td>0.98</td>
</tr>
<tr>
<td>HS S1 Pop</td>
<td></td>
<td>0.65</td>
<td>0.95</td>
</tr>
<tr>
<td>S1 S2 Pop</td>
<td></td>
<td>0.72</td>
<td>0.94</td>
</tr>
<tr>
<td>S2 S2 Pop</td>
<td></td>
<td>0.92</td>
<td>0.98</td>
</tr>
<tr>
<td>HS S2 Pop</td>
<td></td>
<td>0.51</td>
<td>0.95</td>
</tr>
</tbody>
</table>

† Numerator of the heritability equation calculated as the covariance between response units and selection units using the full genetic model proposed by Harris (1964).
‡ Numerator of the heritability equation calculated as the covariance between response units and selection units using only the additive variance.
§ Numerator of the heritability equation calculated as the variance among families using the full genetic model proposed by Harris (1964).
Chapter 5. The Contribution of Additive and Dominance Effects to
Inbreeding Depression

A paper to be submitted for publication in Genetics
Brandon M. Wardyn, Jode W. Edwards and Kendall R. Lamkey

Abstract

Inbreeding is common in natural populations and is often used in artificial selection programs as an aid to selection. Given the widespread presence and usage of inbreeding, it is of interest to quantify the genetic effects causing inbreeding depression. We have defined inbreeding depression strictly as a change in the population mean associated with inbreeding. Since a population of individuals are needed to accurately define inbreeding depression it would be useful to quantify the effects of inbreeding depression on an inbred individual. This by definition is not inbreeding depression, but rather a change in performance due to inbreeding. We have quantified this change in performance for 200 individuals in two generations. Breeding values had a much larger correlation with genotypic values than dominance deviations in both the inbred and noninbred individuals, which indicated most of the genetic value of an individual was under additive control. This result was in spite of this reference population containing more dominance variance than additive variance. The correlation between inbred and outbred performance was 0.62 which was shown to be dependent upon the genetic covariance parameters of the population. We concluded that per se selection on inbred genotypic values would result in higher outbred performance of the population; however, this increase would not be as large as direct selection for outbred performance.
Introduction

Inbreeding depression (ID) is loosely defined as a decline in fitness due the increased probability of alleles being identical by descent (Lynch and Walsh, 1988). Hallauer and Miranda (1988 pp. 314-315) outline ID estimates in maize (Zea mays L.) for eight studies (Levings et al., 1967; Sing et al., 1967; Genter, 1971; Harris et al., 1972; Hallauer and Sears, 1973; Cornelius and Dudley, 1974; Rice and Dudley, 1974; Good and Hallauer, 1977) which vary in their method of inbreeding and reference population. Inbreeding depression was consistently observed across all studies for all grain and plant traits with the exception of days to anthesis. Good and Hallauer (1977) compared three methods of inbreeding (selfing, full-sibbing, and full-sibbing followed by selfing) and found no differences between the methods of inbreeding for plant height which agreed with the results reported by Cornelius and Dudley (1974). The only significant difference between the methods for grain yield occurred at near complete homozygosity when selfing was compared to full-sibbing followed by selfing.

All of these studies estimated ID as the difference between the mean of noninbred individuals (F=0) and the mean of inbred (F>0) individuals. Estimated in this way, ID is a population specific parameter. The genetic expectation of the inbred population is the product of the homozygous genotypic value and associated gene frequencies summed over genotypes, while the noninbred mean is defined as the product of the genotypic value and associated Hardy-Wienberg frequencies summed over genotypes (Kempthorne, 1957). To measure ID, a minimum of two generations with different levels of inbreeding are needed. The resulting estimate of ID has an expected value of:
\[ E[\text{ID}] = \sum p_i \delta_i, \] where \( \delta_i \) represents a homozygous dominance deviation and \( p_i \) is the frequency of the \( i \)th allele. Furthermore, an adequate sample of individuals in both the inbred and noninbred state is needed to sample enough genotypes to obtain accurate estimates. This is easily accomplished in most plant species, and numerous ID studies in maize have taken this approach.

Inbreeding depression studies can be classified broadly into studies that estimate the amount of ID and those that study the genetic cause of ID. Estimation type studies of ID have focused on ID at the population or subpopulation level within a species (Hallauer and Miranda, 1988; Husband and Schemske, 1996; Keller and Waller, 2002). These classical studies have traditionally assigned an ID estimate to a sub-population at some biologically meaningful organizational level. The second major focus of ID studies is the genetic cause of ID. It is commonly accepted that ID is a consequence of genetic changes associated with inbreeding. ID at the single locus level is simply the consequence of an interaction (Crow and Kimura, 1996), or more specifically the consequence of directional dominance (Falconer and Mackay, 1996). Initial theory regarded overdominance as the cause of ID. The debate between the predominant type of gene action in maize, dominance or overdominance, has been fiercely debated among maize breeders for a number of years (Crow, 2000). It is currently accepted that dominance is the predominant type of gene action in maize and overdominance, while most likely present, does not describe the majority of loci in terms of gene action. Recent work attributes ID to deleterious mutations (Charlesworth and Charlesworth, 1987; Carr and Dudash, 2003), which under the dominance hypothesis, should be purged from the population with inbreeding.
ID is influenced not only by the interaction within a locus, but also by interactions among loci (i.e. epistasis). One classical approach for detecting epistasis is to test for a quadratic response in the regression of trait value on the inbreeding coefficient. Departures from linearity are expected only under models with epistasis. The evidence in the literature is mixed on the importance of epistasis to ID and appears to vary not only by species but by trait and genetic background within species as well (Koelewijn, 1998; Dudash et al., 1997; Carr and Dudash, 1997; Willis, 1993). If epistasis is present and the base population is in gametic phase equilibrium, only epistatic interactions involving dominance contribute to ID (Anderson and Kempthorne, 1954; Bulmer, 1980; Hill, 1982; Lynch, 1991). As pointed out by Lynch and Walsh (1998), linearity should not be taken as evidence for the absence of epistasis as positive and negative epistatic effects may cancel out and result in linearity. They also argue that departures from linearity would be rare given the small contributions of dominance x dominance and higher order forms of epistasis relative to the contributions from single locus dominance effects to genotypic value.

**Definition of a reference population**

Before investigating the causes or amount of ID, it is crucial that a precise definition of ID be defined. The formal definition of ID is a change in the mean of a reference population associated with inbreeding. This definition requires that a definition of the reference population should accompany any ID estimate. Consider a population where all individuals are contained within a base population where $F = 0$, ID would be estimated as the difference between the noninbred base population and the same population with $F > 0$. If the reference population was a subpopulation, ID would be estimated as the difference between the mean of the noninbred subpopulation and the mean of the inbred subpopulation. The
reference population can also be defined as an individual, and ID would be estimated as the difference between the noninbred individual and the mean of its inbred offspring. Regardless of the base population, it is crucial to adequately sample the inbred genotypic array of the inbred progeny as inadequate sampling will bias the ID estimate. It follows that ID cannot be estimated from one inbred progeny. The requirement of adequately sampling an individual’s genotypic array introduces a significant obstacle if one wants to assign an ID estimate to an individual as numerous inbred progeny must be developed and evaluated. Another complication that arises when estimating ID of an individual is that a measurement must be obtained on the individual itself. This may not be feasible for all traits and accurate estimates prove to be elusive without the ability to clone the founding individual for replicated testing.

Since ID is defined with respect to a particular reference population, the ID estimate of the reference population cannot be assigned to sub-units of the reference population. Every individual in a population will have its own, unique ID estimate, as alluded to earlier however, the problem with estimating ID for an individual is that a large number of inbred progeny have to be created and evaluated. In most artificial selection programs, the number of progeny from an individual is either limited due to the biology of the organism or the cost associated with large scale production and evaluation of the inbred lines. Most often, only a handful of individuals are used in a breeding program and from these individuals, only a few inbred progeny are developed. The same scenario is found in many natural populations. For many organisms, the number of progeny produced is limited due to the reproductive capacity of the respective organism. For these species, it is impossible to sample enough progeny to encompass the genotypic array of the parent. It may be possible to sample the genotypic array of an individual for traits influenced by one or two genes, although this requires the
assumption that all genes influencing a trait are known which may prove to be a more difficult assumption to satisfy than the sampling dilemma.

In artificial or natural populations, a true estimate of ID of the individual is not as important as the genetic cause of the change in performance of an inbred individual. If for a moment we consider only completely inbred individuals and let each individual’s genetic value (G) be composed of two parts: additive effects (A) and dominance effects (D), it would be informative to know which effect is causing the change in performance (we are assuming epistasis is absent). If the change in performance is due solely to changes in A, inbred per se performance will be perfectly correlated with the associated outbred performance of the respective inbred individual. Thus only inbred per se values are needed to correctly rank the inbred lines in terms of outbred performance. If change in performance is due solely to changes in D, inbred per se performance will give no indication as to the relative outbred performance of the inbred individual. Most likely, both effects are involved in change in performance to some degree. Thus, it would be useful to know the relative influence of the additive effect and dominance effect with regards to the performance of an inbred individual relative to its noninbred parent.

From a selection standpoint, the probability of superior individuals contributing to the next generation will depend on per se performance of the parents. The performance of the next generation relies on the relative effects of A and D to per se performance. If inbreeding is prevalent and the associated change in performance is due solely to a loss of favorable additive effects, high performing individuals will produce high performing inbred and outbred progeny. If change in performance is due solely to dominance effects, the relationship between per se and noninbred progeny performance is not as clear. Thus we are
asking several questions with this research: i) Is there variation for the ID/change in performance among individuals? ii) What is the distribution of dominance deviations? iii) Are additive effects or dominance effects more influential to change in performance? iv) Is it possible to develop high performing noninbred individuals that are also high performing in an inbred state? v) Does the genetic variance structure of a population influence the relationship between inbred and outbred performance?
Materials and Methods

Genetic Model

Harris (1964) defines the genetic value of an individual as:

\[ g_{ij} = \mu + \alpha_i + \alpha_j + \delta_{ij} \]

where:

- \( g_{ij} \) = the genetic value of genotype \( A_iA_j \)
- \( \mu \) = the population mean
- \( \alpha_i \) = the additive effect of allele \( A_i \)
- \( \delta_{ij} \) = the dominance deviation of genotype \( A_iA_j \)

Given the lack of consistent data for epistasis and the small contribution to ID as outlined by Lynch and Walsh (1988), we proceeded with the assumption that epistasis is absent; with a caveat being that epistasis is likely present but to an unknown degree.

As diagramed in the model, once an individual’s genotypic value (G) is deviated from the mean, it is comprised of only two components: additive effects (A) and dominance effects (D). As suggested by Harris (1964), the values of G, A, and D are defined with respect to a panmictic reference population and are thus independent of inbreeding. Thus for each individual in the experiment, G, A, and D values are estimable. The expected value of G at panmixia is zero whereas with inbreeding the expected value of G is a function of ID:

\[ F \sum_i p_i \delta_{ii} \], where \( \sum_i p_i \delta_{ii} \) is the expected amount of ID.

Reference Population

The maize (\textit{Zea mays} L.) population Iowa Corn Borer Synthetic No. 1 (BSCB1), a member of the non stiff stalk heterotic pattern, was the reference population in this study. BSCB1 was developed in the 1940’s at Iowa State University via an intermating of 12 inbred
lines (Penny and Eberhart, 1971). See Hagdorn et al. (2003) for an outline of the 12 parents. BSCB1 was developed via reciprocal recurrent selection (Comstock et al., 1949) with Iowa Stiff Stalk Synthetic. Details for the first five cycles of selection in BSCB1(R) can be found in Penny and Eberhart (1971). Please refer to Holthaus and Lamkey (1995), Keeratinijakal and Lamkey (1993), and Schnicker and Lamkey (1993) for details concerning the selection and breeding methods for cycles six through 11. BSCB1 is considered an elite maize germplasm source that has been the source of elite inbreds for a number of years. For the current study, cycle 13 of BSCB1 was the germplasm source.

**Mating Design**

The mating design used in this study was developed with the intention of estimating the five genetic covariance parameters (Cockerham, 1983). Details concerning formation of the breeding design can be found in Wardyn et al. (2006). We included generations that represented individuals in the S0, S2, and S4 generations as well as the outbred progeny of the S0 and S4 individuals. Inclusion of the outbred progeny of the S0 and S4 individuals allowed us to estimate G, A, and D for each of the S0 and S4 individuals. The S1-TC and S5-TC generations are the equivalent of half-sibs for each S0 and S4 individual. By doubling each half-sib G we were able to obtain a clean estimate of D in each generation.

**Field Procedures**

Four hundred random plants were self-pollinated in cycle 13 of the BSCB1 population. Each S1 ear was planted ear-to-row and the first three plants in the row were self-pollinated via hand pollination. One ear was randomly chosen and the resulting S2 seed was planted the following year and the process was repeated until S5 seed was obtained. Thus, each S0 plant was represented in five generations of inbreeding where each generation
was a direct descendant of the initial S0 plant. The only selection applied during the inbreeding process was that enough seed be present to plant a full nursery row the following year, which in essence was very low intensity selection for fecundity. Seed quantities for the S1 and S5 generations were increased via sib-mating within a nursery row. Effort was made to use each plant once as either a male or a female and reciprocal crosses were not made. A balanced bulk of approximately 10 ears from each nursery row was made and used as the source for yield trial plots.

In addition to the nursery rows, each S1 and S5 line was planted in isolation with BSCB1(R)C13 (the base population from which the 200 individuals were randomly chosen). The S1 and S5 lines were detasseled and used as females, being open-pollinated with BSCB1(R)C13 used as the male. All plants in a row were harvested and shelled in bulk. An individual yield trial plot consisted of two rows, both 5.49 m in length with 0.76 m between rows. Seeds were machine planted at a density of 76 540 plants ha\(^{-1}\) and thinned to 62 190 plants ha\(^{-1}\). Data were taken on an individual plot basis and collected on harvestable grain weight (g adjusted to 15.5 % grain moisture).

**Experimental Design**

The 1 000 entries (200 lines in each of five generations) were planted near Ames, Carroll, Crawfordsville, and Rippey, IA, in 2004 and near Ames and Carroll, IA in 2005. Each location contained two replications of the experiment, laid out in a split-plot design. The 1 000 entries were blocked according to generation (whole plots) due to differences in vigor. The whole plots were laid out in a randomized complete block design. Sub-plots consisted of the 200 entries which were randomized in a 10 by 20 row-column alpha lattice.
The 200 entries were randomly selected from approximately 350 entries with enough remnant seed to plant the necessary nursery and isolation rows for seed increase.

**Data Analysis**

Grain yield was the trait analyzed. Environments (location-year combinations), replications within environments, and lattice columns were treated as fixed effects. Environments were not of primary interest in this study and were thus treated as fixed effects. ID was calculated from the generation means of the S1, S3, S5, and S1-TC generations where the S1-TC generation was considered to have $F=0$ and all generations and their interaction with environments were considered fixed effects. The generation means were then regressed onto $F$ and ID at $F=1$ determined. Both linear and quadratic effects were initially fit and model selection was based on a significance level of 0.05. Individual line ID rates calculated for each individual line via a mixed model where environments, reps within environment, and lattice columns were fitted as fixed effects. Generations, lines and the interaction between lines and the inbreeding level ($F$) were considered random. The variance-covariance matrix for generations was specified to have a block diagonal structure with identical blocks. Best linear unbiased predictors (BLUPs) were predicted for each level of the line by $F$ interaction which represented the rate of ID for each individual line.

Genotypic values ($G$), additive effects ($A$), and dominance deviations ($D$) were estimated via a mixed linear model. Environments, replications nested within environments, and lattice columns were considered fixed effects. In addition, $F$ was nested within environments and was also considered fixed. The nested $F$'s were the rates of ID for each environment. Generations were considered random with a block diagonal covariance matrix of identical blocks. Each block represented one of the 200 lines with an unstructured
variance-covariance matrix. BLUPs were obtained for each line by generation combination, which resulted in 800 BLUPs (200 lines in four generations). ID effects for the inbred generations were calculated for the S1 and S5 generations via a linear function of the individual environment ID effects. G values for each line in the S1 and S5 generation were obtained by adding the BLUPs to the respective ID effects. A values were obtained from the line by generation BLUPs for the S1-TC and S5-TC generations. For example, the A value for line 23 in the S1 generation was calculated as: 2 * BLUP estimate from the S1-TC generation for line 23. D values were the difference between G and A within the same generation. Thus we had G, A, and D values for each of the 200 lines in both the S0 and S4 generations. The S3 generation was left out of this part of the analysis since we did not have estimates of the outbred progeny.
Results and Discussion

Inbreeding Depression

It has been well documented in the literature that ID exists for nearly all traits in maize. We did find significant base population ID for grain yield with no significant deviation from linearity, and concluded that if epistasis is influencing ID it is doing so at a small level (Table 10). Since we did not evaluate multiple inbred lines from the same founder individual, we were unable to calculate individual ID estimates. We could, however, calculate change in performance estimates for each of the 200 lines as we did have estimates of noninbred performance and measures of inbred performance for a single individual derived from each founder. Significant variation for change in performance did exist as change in performance ranged from roughly -1.5 to -6.5 mg ha-1 when expressed as a difference between noninbred and completely inbred individuals (Fig. 4). Also of note is that all individuals experienced a negative change in performance, which was expected given the large amounts of ID normally observed in maize for grain yield.

Although our design did not allow us to estimate ID at the individual level, we could estimate D at the individual level. At this point it is crucial that we emphasize that D is an estimator of $\delta$, A is an estimator of $\alpha_i + \alpha_j$, and G is an estimator of the sum of A and D for an individual. By evaluating only one inbred individual, the expected values of G, A, and D are dependent upon the allelic content of the inbred individual such that the expected values are defined as a conditional expectation depending on the genotype of the parent. For inbred individuals, the conditional expectations of G, A, and D, when deviated from the population mean are:
\[ E(A | A_i A_i) = 2\alpha_i \]
\[ E(D | A_i A_i) = \delta_{ij} \]
\[ E(G | A_i A_i) = 2\alpha_i + \delta_{ij}. \]

For noninbred individuals, the conditional expectations, when deviated from the population mean are:

\[ E(A | A_i A_i) = \alpha_i + \alpha_j \]
\[ E(D | A_i A_i) = \delta_{ij} \]
\[ E(G | A_i A_i) = \alpha_i + \alpha_j + \delta_{ij}. \]

Our mating design took advantage of these conditional expectations in that we were able to estimate G, A, and D for each of the 200 individuals in the experiment in both the noninbred and inbred state. The S1 family, when taken as a group, represented the noninbred S0 parent. Thus, the conditional expectations for G, A, and D can be calculated from the S1 and S1-TC generations:

\[ E(G) = E(S_{1\text{mean}}) = \alpha_i + \alpha_j + \frac{1}{2} \delta_{ij} + \frac{1}{4} (\delta_{ii} + \delta_{jj}) \]
\[ E(A) = 2 * E(S_{1-TC \text{mean}}) = \alpha_i + \alpha_j \]
\[ E(D) = [E(G) - 2 * E(A)] = \frac{1}{2} \delta_{ij} + \frac{1}{4} (\delta_{ii} + \delta_{jj}). \]

Since we are taking the mean of the S1 family, these expectations represent the founding noninbred S0 individual. Although our design did not contain completely inbred individuals, the S5 generation represented the S4 individual with an F value of 0.9375, which we considered close enough to complete homozygosity that D estimates will be representative of D estimates of completely inbred individuals. Thus we will continue under the assumption that any increase in homozygosity beyond 0.9375 will not change our estimates to a large
degree and with the realization that our D estimates actually follow a mixture distribution (Edwards, 2006). The conditional expectations of G, A, and D for an inbred individual in our mating design, when deviated from the population mean are:

\[ E(G) = E(S_{5, mean}) = 2\alpha_i + \delta_{ii} \]
\[ E(A) = 2 \cdot E(S_{5} - TC_{mean}) = 2\alpha_i \]
\[ E(D) = \left[ E(G) - 2 \cdot E(A) \right] = \delta_{ii}. \]

In our mating design, both inbred and outbred D values are estimable. These D values are estimators of each individual’s dominance deviation. Estimates of D obtained from the S1 generation were estimators of the quantity: \( \frac{1}{2} \delta_{ij} + \frac{1}{4} (\delta_{ii} + \delta_{jj}) \), whereas D estimates from the S5 generation were estimates of \( \delta_{ii} \). When taken as a group, our D values from the S5 generation estimate the average ID for all the 200 individuals i.e. the ID of the population. The following equations outline this relationship:

\[ E[ID] = \sum p_i \delta_{ii} \]
\[ E[D] = \delta_{ii} \quad \text{when D is estimated from inbred individuals.} \]

Our D estimates were normally distributed in both the S1 and S5 generations (Fig. 5). The S1 generation has a less negative mean and smaller variance than the S5 generation. This fits expectations because the S5 generation contains more homozygous dominance deviations than the S1 generation due to the higher level of inbreeding, and given the prevalence of ID in maize, this results in a lower mean. The larger variance in the S5 generation is also not surprising since homozygous dominance deviations need to be larger than heterozygous dominance deviations for ID to exist. These results also indicate that the variation due to dominance is more pronounced in inbred lines.
Genetic cause of change in performance

Based solely on additive expectations, the variation among noninbred individuals should be equal to half the variation among inbred individuals. A completely additive model is inadequate to describe all of the variation in most species as dominance will also contribute to the variation among individuals. In this maize population, the estimate of the dominance variance is larger than the estimate of the additive variance for grain yield; as the ratio of dominance to additive variance is 1.34 (Wardyn et al., 2006). If dominance variance constitutes such a large part of the total genetic variance, it is of interest to know how much of the variation among individuals is due to D relative to A. Since selection is practiced on the level of the individual, any selection that happens before reproduction will be based on per se performance. In a natural population, selection on per se performance has the potential to have detrimental effects on the population performance. If inbreeding is prevalent, individuals may be deemed superior due to their lack of highly negative D values. If these individuals also contain below average A values, the future performance of the population will be compromised. The same scenario can be developed in an artificial selection program where parents are selected solely on per se performance. Quantification of the relative influence of A and D to G will give insight as to which effect is influencing per se performance.

Since we have estimates of G, A, and D for a large collection of individuals, we can quantify the relative impact of A and D on G. In the S1 generation, the correlation between G and A is high (corr. = 0.83) whereas the correlation between G and D is much lower (corr. = 0.52) (Fig. 6). This indicates that in noninbred individuals, additive effects have a larger influence on G than dominance effects. The same is also true when inbred individuals are
analyzed (Fig. 7). It does appear, however, that as individuals become more inbred, A is a less reliable predictor of G as the correlation drops from 0.83 to 0.71. This reduction is accompanied by a similar reduction in the correlation between D and G. Given the large amount of dominance variance in this population, one would expect D to have an equal or larger impact on G than A. Our results indicate, however, that G is best described by A in both noninbred and inbred individuals. It appears that in this population, selection on per se values will be acting predominantly on A given the high correlation between G and A. The correlation is not perfect, however, which indicates that selection on outbred performance will be slightly more effective at identifying high A values than per se selection on the inbred individuals. The drawback to evaluating outbred progeny is that another generation is required to form the outbred progenies.

Up to this point we have been describing A and D as independent effects, which is not entirely correct. Since both A and D are intra-locus effects and are determined by the same two alleles (in a diploid organism), their respective effects are fixed in a homozygous individual. The result of this is that the homozygous genotype which contains the largest A will be associated with a particular D value. In this aspect, an inbred individual's A value is associated with specific value of D. What we are concerned with, however, is the relative values of the associated A and D values within individuals. Specifically, we need to know if high A values are associated large or small D values. Given that D values are negative, an ideal situation would be one where high A values are associated with less negative D values. By definition, A and D values are uncorrelated in noninbred individuals (Fig. 8). Thus in noninbred individuals, A and D act as independent effects due to the fact that D is a result of an intra-locus interaction. This independence does not hold, however, for inbred individuals.
We found that with an F value of 0.9375, A and D values have a correlation of -0.37. From this we can conclude that within an inbred individual, a high A value is more likely to be associated with a large negative D than with a small negative D. This correlation is a result of a negative covariance between homozygous dominance deviations and additive effects. Harris (1964) assigns the parameter D1 to this covariance. Previous estimates of D1 have all shown that this covariance tends to be negative (Wardyn et al., 2006; Edwards and Lamkey, 2002; Coors, 1988) which agrees with our current findings.

What is not known is the cause of a large and negative D1 estimate. Although BSCB1 has been predominately under selection for grain yield, it has also been under selection for grain moisture. Both grain moisture and grain yield contain negative estimates of D1, but the magnitude of D1 varies. D1 is 19% as large as the sum of additive and dominance variance for grain yield, but only 12% as large as the same sum for grain moisture (Wardyn et al., 2006). It appears that the relative proportions of additive and dominance variance influences D1; as the dominance to additive variance ratio decreases the magnitude of D1 also decreases. It has been hypothesized that repulsion phase linkages brought about by selection may have caused the large negative estimate of D1 in this population (Edwards and Lamkey, 2002). While we can neither prove nor disprove this hypothesis with this data, it is a plausible explanation given the high selection intensity and finite population size in BSCB1. The hypothesis does not explain, however, why grain moisture failed to develop the same repulsion phase linkages given that it was also under selection.

**Inbred vs. Outbred Performance**

In noninbred individuals, A and D are independent and thus uncorrelated; however in inbred individuals a correlation is established as loci are fixed and additive effects become
associated with specific homozygous dominance deviations. The phenotypic consequence of this correlation is that inbred per se performance is a function of not only A, but also the associated homozygous dominance deviation. As F increases, the A of each locus is simultaneously being fixed. This A contributes equally to the outbred and inbred performance since all genotypic values are deviated from the panmictic mean. The other consequence of inbreeding, ID, only manifests itself in the inbred individual in the form of a homozygous dominance deviation. Thus, the quantitative genetic reasons for the discrepancy between inbred and outbred performance is homozygous dominance deviations and their covariance with breeding values. The following equation outlines the relationship between inbred performance and outbred performance within a population:

\[
\text{Corr}(X_{\text{inbred}}, X_{\text{outbred}}) = \frac{\sigma^2_A + D_1}{\sqrt{2\sigma^2_A + 4D_1 + D_2^*}} \frac{1}{\sqrt{2\sigma^2_A}}
\]

where,

\(X_{\text{inbred}}\) = the phenotypic value of a completely inbred individual (X) and
\(X_{\text{outbred}}\) = the phenotypic value of the half-sib family produced from the cross of individual X to the source population
\(D_1\) = the covariance between additive effects and homozygous dominance deviations
\(D_2^*\) = the variance of homozygous dominance deviations

(Cockerham, 1983; Cockerham and Matzinger, 1985).

This equation is an expected value for all individuals and it demonstrates that the relationship between inbred and outbred performance is dependent upon parameters relating to the homozygous dominance deviations: D1 and D2*. In the absence of dominance, it is obvious that the correlation between inbred and outbred performance is one. Based on previous genetic covariance parameter estimates in this population (Wardyn et al., 1996), the
correlation between inbred and outbred performance is 0.62 and thus it appears plausible that high performing inbred individuals will produce high performing outbred progeny. Genetic covariance parameter estimates have been obtained for another maize population, BS13, with a slightly different genetic structure as well. The estimates of additive and dominance variance were nearly equal in BS13, but D1 was proportionately larger than in BSCB1 (Edwards and Lamkey, 2002). The correlation between inbred and outbred performance was only 0.34 in this population which illustrates the importance of the underlying genetic structure on inbred and outbred performance.

Since the correlation between inbred and outbred performance is positive, we can conclude that additive effects, i.e. A values, are the predominant cause change in performance in this population. The correlation acts as a scale to which the effects, A or D, influence change in performance. A correlation of one indicates complete additive control, negative one indicates complete dominant control, and zero indicates equal contributions from additive and dominance effects. It is often assumed that if family level ID exists, then variation among the inbred progeny of the respective families is mainly a function of the family level estimate of ID and hence due to D. With respect to the base population, however, this variation cannot all be attributed to ID. This work demonstrates that although the classical definition of ID is only a function of D, the decline in performance often attributed to ID can be caused by changes in A in addition to D. It also demonstrates the importance of measuring ID correctly, as many measurements of ID are confounded by A (Fox, 2005).

Future studies on this topic are needed in populations not affected by artificial selection. We are unable to prove that the genetic relationships we found are natural
occurrences or if they have been brought upon by artificial selection. We know that the underlying genetic architecture of a trait influences the relationship between inbred and outbred performance, but what is unknown is how selection is affecting the underlying genetic structure of a population. From the evidence found in maize, it appears that the loss of favorable additive effects has more influence than dominance on the change in performance of an individual associated with inbreeding.
References


Willis, J. H. Effects of different levels of inbreeding on fitness components in Mimulus guttatus. Evolution 47:864-876.
Table 10. Inbreeding depression estimates for five individual environments and combined analysis over all five environments for grain yield.

<table>
<thead>
<tr>
<th>Location</th>
<th>Inbreeding Depression</th>
<th>Grain Yield (Mg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ames-2004</td>
<td>-2.57** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Crawfordsville</td>
<td>-4.27** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Carroll-2004</td>
<td>-2.77** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Rippey</td>
<td>-4.95** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Ames-2005</td>
<td>-4.13** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Carroll-2005</td>
<td>-5.43** ± 0.39</td>
<td></td>
</tr>
<tr>
<td>Combined†</td>
<td>-4.02** ± 0.37</td>
<td></td>
</tr>
</tbody>
</table>

† Inbreeding depression is defined as the difference between the noninbred (F=0) and inbred (F=1) generations.
‡ Combined analysis across all locations.
Figure Captions

Figure 4. The distribution of the change in performance values from F=0 to F=1 for 200 individuals.

Figure 5. The distribution of dominance deviations in the S1 and S5 generations.

Figure 6. Scatter plots of the dominance deviation and breeding value estimates plotted against the genotypic values for 200 individuals in the S1 generation.

Figure 7. Scatter plots of the dominance deviation and breeding value estimates plotted against the genotypic values for 200 individuals in the S5 generation.

Figure 8. Scatter plots of the dominance deviation and breeding value estimates for 200 individuals in the S1 and S5 generation.
Figure 4

![Histogram showing change in performance (CP) with mean -4.27, variance 1.09.](image)
Figure 5

**S1 GENERATION (mg/ha)**

- **Mean**: -2.21
- **Variance**: 0.17

**S5 GENERATION (mg/ha)**

- **Mean**: -4.43
- **Variance**: 0.57
Figure 6

GRAIN YIELD (mg/ha): S1

BREEDING VALUE

GENOTYPIC VALUE

corr = 0.83

GRAIN YIELD (mg/ha): S1

DOMINANCE DEV

GENOTYPIC VALUE

corr = 0.52
Figure 7

**GRAIN YIELD (mg/ha): S5**

![BREEDING VALUE vs. GENOTYPIC VALUE](image1)

- Correlation: 0.71

---

**DOMINANCE DEV vs. GENOTYPIC VALUE**

- Correlation: 0.39
Figure 8

**GRAIN YIELD (mg/ha): S1**

![Graph showing the relationship between breeding value and dominance deviation for GRAIN YIELD (mg/ha): S1.](image)

**GRAIN YIELD (mg/ha): S5**

![Graph showing the relationship between breeding value and dominance deviation for GRAIN YIELD (mg/ha): S5.](image)