Genomewide Selection: Prediction Accuracy, Marker Imputation, and Introgression of Semidwarf Corn Germplasm

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Emily Elizabeth Combs

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Rex Bernardo

December 2012

© Emily Elizabeth Combs 2012

Acknowledgements

I am profoundly grateful to my advisor, Dr. Rex Bernardo, for his teaching, mentorship and guidance throughout my time at the University of Minnesota. I deeply appreciate the opportunity to learn and work with him. I am also thankful to my advisory committee, Drs. Yang Da, Craig Sheaffer, Kevin Smith and Julia Zhang, who always gave helpful advice and productive discussion. I would also like to thank all of the instructors and other mentors who have helped me during my degree, at the University of Minnesota, CIMMYT and other institutions.

I thank the members of the Corn Breeding and Genetics Project, especially Eric Ristau, for his assistance with my field and greenhouse experiments. I am deeply grateful for the guidance, assistance and friendship of fellow graduate students Amy Jacobson, Lisa Marie Krchov, Lian Lian, Jon Massman, Chris Schaefer, Addie Thompson, and Cathrine Ziyomo. I thank the staff of the Minnesota Agriculture Experiment Stations at Lamberton, Steve Quiring, Waseca, Tom Hoverstad and Matt Bickel, Rosemount, Gerry Holz, and the Plant Growth Facilities, Roger Meissner, Pam Warnke, and Dean Ziertman for their assistance and encouragement during my field research.

I thank the University of Minnesota, Dupont Pioneer and the USDA for providing financial support for graduate studies. I also thank the University of Minnesota Super Computing Institute for computational resources that helped me complete my projects.

I am also thankful for my family and friends who have tirelessly supported me throughout my time in graduate school. I especially thank my parents, Terri and Ed, for their support and encouragement and my siblings Katie, Brian and John.

Abstract

I present here three studies on genomewide selection, a marker based selection procedure with the potential to accelerate genetic gain while decreasing costs. For the first study, I looked at factors that have been previously derived by other researchers as determining the accuracy of genomewide selection: training population size (N), trait heritability (h^2), and effective number of loci or chromosome segments underlying the trait (M_e). My objective was to determine if prediction accuracy is equal across traits if N, h^2 and marker number (N_M) are kept constant. Cross validations indicated that the traits predicted most accurately did not always stay the same across changes to h^2, N, and N_M. For the second study, I investigated the use of marker imputation to reduce costs by genotyping the training population with many markers (M_{Total}), genotyping the validation population with fewer markers (M_{Low}), and predicting the genotypes at the $M_{Total} - M_{Low}$ markers in the validation population. My objective was to determine if genomewide prediction with imputed markers can be as accurate as genomewide prediction with non-imputed markers in inbred collections. With imputation, many combinations of M_{Total} and M_{Low} led to prediction accuracy that was as high as the accuracy with M_{Total} non-imputed markers. For the third study, I used a semidwarf corn (*Zea mays* L.) line that could potentially be grown in new areas of production or in alternative crop rotations. My objectives were to determine: (i) if genomewide selection is useful for the rapid improvement of an adapted × exotic cross; and (ii) if 4 cycles of genomewide selection are more effective than phenotypic backcrossing to the BC_4 for a trait with major genes. Genomewide selection from Cycle 1 until Cycle 5 either maintained or improved upon the gains from

phenotypic selection achieved in Cycle 1. Compared with phenotypic backcrossing, genomewide selection led to better mean performance and a higher proportion of exotic germplasm introgressed. To my knowledge, this is the first empirical study on genomewide selection to improve an exotic × adapted cross.

Table of Contents

List of Tables vi
List of Figures viii
Chapter 1: Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers 1
Introduction 2
Materials and Methods 3
Simulated and Empirical Populations 3
Results and Discussion 8
Easily Controllable Factors: Marker Density and Population Size 8
Influence of Heritability 10
Implications 14
Chapter 2: Marker imputation prior to genomewide prediction among diverse maize inbreds and elite barley inbreds 19
Introduction 20
Materials and Methods 23
Overview 23
Genotypic and phenotypic data 23
Selection of M_{Low} marker sets 25
Linkage disequilibrium 26
Marker imputation 27
Genomewide prediction accuracy 28
Results and Discussion 30
Linkage disequilibrium and imputation accuracy 30
Accuracy of genomewide prediction with and without imputation 31
Chapter 3: Genomewide Selection to Introgress Semidwarf Corn Germplasm into U.S. Corn Belt Inbreds 42
Introduction 44
Materials and methods 45
Germplasm 46
Cycle 0 and Parental Screens 46
Multi-trait Selection Indices 49
Phenotypic Selection of Cycle 0 Individuals to Form Cycle 1 50
Genomewide Selection in Cycles 1 to 5 51
Phenotypic Backcrossing 52
Progress from Selection 52
Predicted Response and Population Genetics of Cycle 0 to Cycle 4 Entries 54
Results and Discussion 55
Performance of Cycle 0 55
Response to Genomewide Selection 56

Correlated Responses with a Different Tester 60
Genomewide Selection vs. Phenotypic Backcrossing 61
Prospects for Developing Semidwarf Hybrids Adapted to the U.S. Corn Belt 63
Bibliography 73
Appendix 79

List of Tables

Chapter 1

Table 1 Number of single nucleotide polymorphism markers, spacing between adjacent markers, and linkage disequilibrium (r^2) for the low, medium, and high density marker sets in each population. 15

Table 2 Heritability (h^2), observed genomewide prediction accuracy (r_{MP}), and predicted r_{MP} assuming different effective numbers of chromosome segments (M_e) for different traits in different populations. 16

Chapter 2

Table 3 Linkage disequilibrium (LD), accuracy of imputation (r_I), and accuracy of genomewide prediction (r_{MP}) with different numbers of single nucleotide polymorphism markers in a collection of 283 maize inbreds. Results are for a single partitioning of the maize inbreds into a training population and test population. 36

Table 4 Linkage disequilibrium (LD), accuracy of imputation (r_I), and accuracy of genomewide prediction (r_{MP}) with different numbers of single nucleotide polymorphism markers in a collection of 860 barley inbreds. Results are for a single partitioning of the barley inbreds into a training population and test population. 38

Table 5 Mean accuracy of imputation (r_I) and genomewide prediction (r_{MP}) across five repeats for selected numbers of single nucleotide polymorphism markers in a collection of 283 maize inbreds. 39

Table 6 Mean accuracy of imputation (r_I) and genomewide prediction (r_{MP}) across five repeats for selected numbers of single nucleotide polymorphism markers in a collection of 860 barley inbreds. 41

Chapter 3

Table 7 Trait means, testcross genetic variance (V_G), entry-mean heritability (h^2) and correlation between predicted and observed performance (r_{MP}) in Cycle 0 testcrosses (to DC2) of 152 (DC1 × LH74) F_3 families and 163 (DC1 × PHG50) F_3 families. 66

Table 8 Testcross performance (with DC2 as the tester) of different cycles of genomewide selection and of progeny from phenotypic backcrossing in the DC1 × LH74 and DC1 × PHG50 corn crosses. 67

Table 9 Changes in single nucleotide polymorphism (SNP) allele frequency, SNP-based inbreeding coefficient, and linkage disequilibrium (mean r^2 between adjacent SNP markers) for each cycle of genomewide selection in the DC1 × LH74 and DC1 × PHG50 corn crosses. 687

Table 10 Testcross performance (with a nondwarf inbred, LH227, as the tester) of different cycles of genomewide selection and of progeny from phenotypic backcrossing in the DC1 × LH74 and DC1 × PHG50 corn crosses. 70

List of Figures

Chapter 1

Figure 1 Accuracy of genomewide prediction (r_{MP}) with different levels of heritabilty (h^2). Results are for the highest marker density and training population size within each population. 18

Chapter 2

Figure 2 Marker imputation scheme. 42

Chapter 3

Figure 3 Plot of QTL mapping results for plant height based on F_3 family testcrosses to DC2. 71

Figure 4 Predicted and observed responses to genomewide selection in the DC1 × LH74 and DC1 × PHG50 corn crosses. A * indicates that the 95% confidence interval on the observed response did not overlap with the predicted response. 71

Chapter 1: Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers

In genomewide selection, the expected correlation between predicted performance and true genotypic value is a function of the training population size (N), heritability (h^2), and effective number of chromosome segments underlying the trait (M_e). My objectives were to (i) determine how the prediction accuracy of different traits responds to changes in N, h^2, and number of markers (N_M), and (ii) determine if prediction accuracy is equal across traits if N, h^2, and N_M are kept constant. In a simulated population and four empirical populations in maize (*Zea mays* L.), barley (*Hordeum vulgare* L.), and wheat (*Triticum aestivum* L.), I added random nongenetic effects to the phenotypic data to reduce h^2 to 0.50, 0.30 and 0.20. As expected, increasing N, h^2, and N_M increased prediction accuracy. For the same trait within the same population, prediction accuracy was constant for different combinations of N and h^2 that led to the same Nh^2. Different traits, however, varied in their prediction accuracy even when N, h^2, and N_M were constant. Yield traits had lower prediction accuracy than other traits despite the constant N, h^2, and N_M. Empirical evidence and experience on the predictability of different traits are needed in designing training populations.

Introduction

Genomewide selection (or genomic selection) allows breeders to select plants based on predicted instead of observed performance. In genomewide selection, effects of markers across the genome are estimated based on phenotypic and marker data from a training population (Meuwissen et al., 2001). The marker effects are then used to predict the genotypic value of individuals that have been genotyped but not phenotyped. The effectiveness of genomewide selection depends on the correlation between the predicted genotypic value and the underlying true genotypic value (Goddard and Hayes, 2007).

The expected accuracy of genomewide selection has been expressed as a function of the size of the training population (N), trait heritability (h^2), and the effective number of quantitative trait loci (QTL) or chromosome segments underlying the trait (M_e; Daetwyler et al., 2008; 2010):

$$r_{g\hat{g}} = \sqrt{\frac{Nh^2}{Nh^2 + M_e}} \quad \text{[Eq. 1]}$$

The M_e refers to the idealized concept of having a number of independent, biallelic, and additive QTL affecting the trait (Daetwyler et al., 2008), and M_e has been proposed as a function of the breeding history of the population and of the size of the genome (Goddard and Hayes, 2009; Hayes and Goddard, 2010; Meuwissen, 2012). Equation 1 also assumes that the number of markers (N_M) is large enough to saturate the genome.

Equation 1 and previous simulation and cross-validation studies have indicated that prediction accuracy generally increases as N increases (Lorenzana and Bernardo, 2009; Grattapaglia and Resende, 2011; Guo et al., 2011; Heffner et al., 2011a; Heffner et al., 2011b; Albrecht et al., 2011), as h^2 increases (Lorenzana and Bernardo, 2009; Guo et al., 2011; Heffner et al., 2011a; Heffner et al., 2011b; Resende et al., 2012), and as the

number of QTL decreases (Zhong et al., 2009; Grattapaglia et al., 2009; Lorenz et al., 2011). However, previous research has focused largely on the effects of N, h^2, and N_M without considering the role that different traits play in determining prediction accuracy. Because traits tend to differ in their h^2, the effects of h^2 in previous empirical studies were confounded with any intrinsic differences in prediction accuracy for different traits. This confounding of h^2 with traits raises the question that if N_M, N, and h^2 are held constant for several traits, would the prediction accuracy be constant across different traits?

By better understanding the factors that affect genomewide prediction accuracy, breeders will be able to design genomewide selection schemes that work best. The objectives of this study were to (i) determine how the prediction accuracy of different traits in plants responds to changes in N, h^2, and N_M and (ii) determine if prediction accuracy is equal across traits if N, h^2 and N_M are kept constant.

Materials and Methods

Simulated and Empirical Populations

I considered five different populations: a simulated biparental population (Bernardo and Yu, 2007); an empirical biparental maize (*Zea mays* L) population (Lewis et al., 2010); an empirical biparental barley (*Hordeum vulgare* L.) population (Hayes et al., 1993); a collection of barley inbreds with mixed ancestry (referred to in the rest of this manuscript as a "mixed population"); and a wheat (*Tritucum aestivum* L.) mixed population. In the simulated population, the genome had 10 chromosomes that comprised 1749 cM (Senior et al., 1996) with N_M = 350 biallelic markers giving a mean marker density of 5 cM. The genome was divided into N_M bins and a marker was located at the

midpoint of each bin. Populations of 300 doubled haploids, developed from a cross between two inbreds, were simulated for a trait controlled by 10, 50, or 100 QTL. The QTL were randomly located across the entire genome. The QTL testcross effects, which are additive (Hallauer and Miranda, 1981), varied according to a geometric series (Lande and Thompson, 1990; Bernardo and Yu, 2007). A maximum h^2 of 0.95 was initially simulated by adding random nongenetic effects drawn from a normal distribution with a mean of zero and the appropriately scaled standard deviation.

The empirical biparental maize population comprised testcrosses of 223 recombinant inbreds derived from the intermated B73 × Mo17 population (Lee et al., 2002). The testcrosses were evaluated in four Minnesota environments in 2007 for grain yield, grain moisture, root lodging, stalk lodging, and plant height (Lewis et al., 2010). Genotypic data for 1339 polymorphic markers covering the approximately 6240 cM linkage map were available from MaizeGDB (Lawrence et al., 2005). By deleting markers with >20% missing data, I retained a maximum of $N_M = 1213$ markers.

The biparental barley population comprised 150 doubled haploids derived from Steptoe × Morex. In 1991, grain yield and plant height were measured in 16 environments; grain protein, malt extract, and alpha amylase activity were measured in nine environments whereas lodging was measured in six environments (Hayes et al., 1993). Genotypic data for 223 polymorphic markers covering the approximately 1250 cM linkage map were available from the USDA-ARS (2008). This number of markers and linkage-map size corresponded to a mean marker density of 5 cM (USDA-ARS, 2008).

The barley mixed population comprised 96 inbreds included in the University of Minnesota barley breeding program preliminary yield trials in 2009. Grain protein, grain

yield, heading date, and plant height were measured in two environments with two replications per environment; data were available as means in each environment. Genotypic data for 1178 polymorphic markers covering the approximately 1250 cM linkage map were available from the Hordeum Toolbox (http://www.hordeumtoolbox.org/). Genotypic and phenotypic data were downloaded from the Hordeum Toolbox on September 2, 2012.

The wheat mixed population comprised 200 inbreds included in a University of Nebraska nitrogen use efficiency trial in 2012. Biomass, heading date, maturity, plant height, and grain yield were measured in two main plots (low N and moderate N) with two replications. For the 200 inbreds genotypic data for 731 polymorphic markers covering the approximately 2,569 cM linkage map (Somers et al., 2004) were available from the Triticeae Toolbox (http://triticeaetoolbox.org/). Genotypic and phenotypic data were downloaded from the Triticaeae Toolbox on October 1, 2012.

Changes in N, N_M, and h^2

I considered 2–3 different training population sizes (N) for each simulated or empirical population. Out of the total number of inbreds (N_{Total}) in each population, I chose N inbreds and considered the $N_V = (N_{Total} - N)$ remaining inbreds as the validation population. I considered the following sizes of the training population: N = 48, 96, and 192 for the simulated population and biparental maize population; N = 48, 72, and 96 for the biparental barley population; N = 72 for the barley mixed population; and N = 72 and 96 for the wheat mixed population.

I considered three different numbers of markers (N_M) for each population (Table 1). To achieve lower marker densities, markers were removed to retain even spacing

between markers. For the wheat mixed population, linkage-map or physical positions were unavailable so markers were removed at random. Higher marker densities were retained in the mixed populations than in the biparental populations because higher coverage levels are needed for accurate predictions in mixed populations than in biparental populations (Lorenz et al., 2011). Due to differences in the types of progeny and structure of the different populations (e.g., doubled haploids versus recombinant inbreds and biparental versus mixed populations), the same marker density in different populations corresponded to different levels of linkage disequilibrium. I therefore calculated the mean pairwise r^2 values between adjacent markers through Haploview (Barrett et al., 2005). This analysis was done for each marker density within each population. Linkage disequilibrium could not be evaluated in the wheat mixed population because of the lack of information on marker positions.

The h^2 of a given trait was left unchanged (i.e., as simulated or as calculated from the data) or reduced to 0.50, 0.30, or 0.20. The h^2 is technically undefined in a collection of inbreds that are not members of the same random mating population. For the mixed-populations, I defined Θ as $\Sigma\tau_i^2/(N - 1)$, where τ_i was the effect of the ith inbred. The ratio between Θ and the total phenotypic variance indicates how much of the observed variation is due to genetic causes. I calculated this ratio, which I refer to as g^2, for each trait in the barley and wheat mixed populations using a mixed model where inbreds had fixed effects and other effects were random.

Reductions in h^2 or g^2 were obtained in a three-step process. First, analysis of variance was conducted on the set of N lines to estimate genetic and nongenetic variance components or Θ. Tests of significance of the genetic variance component or Θ were

conducted and confidence intervals on h^2 or g^2 were constructed (Knapp, 1985). Second, the amount of nongenetic variance required (V_{Extra}) to adjust the observed h^2 or g^2 to the target h^2 or g^2 was calculated. Third, random nongenetic effects were added to the data. These random nongenetic effects were normally and independently distributed with a mean of zero and a standard deviation equal to the square root of V_{Extra}.

Genomewide Prediction and Cross Validation

For the N inbreds in the training population, genomewide marker effects were obtained by ridge-regression best linear unbiased prediction (RR-BLUP) as implemented in the R package rrBLUP version 3.8 (Endelman, 2011) for R version 2.12.2 for Windows 7. The performance of each of the N_V inbreds in the validation set was then predicted as $\mathbf{\hat{y}_p} = \mathbf{M\hat{g}}$, where $\mathbf{\hat{y}_p}$ was an $N_V \times 1$ vector of predicted trait values for the inbreds in the validation set; $\mathbf{M}$ was an $N_V \times N_M$ matrix of genotype indicators (1 and –1 for the homozygotes and 0 for a heterozygote) for the validation set; and $\mathbf{\hat{g}}$ was an $N_M \times 1$ vector of RR-BLUP marker effects (Meuwissen et al., 2001). The accuracy of genomewide prediction was calculated as the correlation (r_{MP}) between $\mathbf{\hat{y}_p}$ and the observed performance of the N_V inbreds in the validation set.

The partitioning of each population into training and validation sets was repeated 500 times, and the prediction accuracies I report were the mean r_{MP} across the 500 repeats. Each repeat comprised a random set of N inbreds and a different set of nongenetic effects used to adjust h^2 or g^2. However, for a given marker density in a population, I used the same set or subset of markers because the subset of markers was chosen to achieve as even spacing as possible between adjacent markers. Least significant differences (LSD, $P = 0.05$) for r_{MP} were calculated for each population using

SAS PROC GLM of the SAS software version 9.2 for Windows 7 (Cary, NC), with the combinations of N, h^2, and N_M as the independent variables.

I also tested combinations of N and h^2 (or g^2) that led to a constant Nh^2 (or Ng^2); for simplicity, the maximum N_M was used. For the simulated population and biparental maize population, I compared r_{MP} with $N = 72$ and $h^2 = 0.50$ ($Nh^2 = 36$) versus r_{MP} with $N = 180$ and $h^2 = 0.20$ ($Nh^2 = 36$). For the biparental barley population and the mixed wheat population, I compared r_{MP} with $N = 72$ and h^2 or $g^2 = 0.50$ (Nh^2 or $Ng^2 = 36$) versus r_{MP} with $N = 120$ and h^2 or $g^2 = 0.30$ (Nh^2 or $Ng^2 = 36$). The same procedures for genomewide prediction and cross validation as described above were used, and the LSD was calculated between the pairs of r_{MP} values.

I also calculated expected prediction accuracy based on Eq. 1 (Daetwyler et al., 2008; 2010) for the largest values of N, h^2, and N_M. Given that r_{MP} was the correlation between predicted genotypic values and phenotypic values, I multiplied $r_{g\hat{g}}$ by h so that the expected prediction accuracy can be directly compared with r_{MP}. Three different values of M_e were used: (i) the number of chromosomes; (ii) the size of the linkage map divided by 50 (i.e., with 50 cM between unlinked loci); and (iii) N_M.

Results and Discussion

Easily Controllable Factors: Marker Density and Population Size

The number of markers (N_M) and size of the training population (N) are the factors that are most easily controlled by the investigator. The accuracy of genomewide predictions (r_{MP}) increased as the number of markers (N_M) increased (Suppl. Tables 1–5). However, gains in r_{MP} began to plateau once a moderately high marker density was reached. This result was important because the expected prediction accuracy (Eq. 1)

derived by Daetwyler et al. (2008; 2010) assumes that the genome is sufficiently saturated with markers, and I surmise that a lack of increase in r_{MP} after a certain N_M is reached indicated marker saturation in the populations I studied. In the biparental populations, there was no consistent gain in r_{MP} from increasing marker density above one marker per 12.5 cM (Suppl. Tables 1, 2, and 5). This result was consistent with the results from QTL mapping in biparental populations, for which markers spaced 10–15 cM apart have been found sufficient (Doerge et al., 1994). The mixed populations generally showed nonsignificant gains in r_{MP} from the moderate marker density (markers spaced 2 cM apart in barley and 4.5 cM apart in wheat) to high density (markers spaced 1 cM apart in barley or 3.5 cM apart in wheat) (Suppl. Tables 3 and 4).

Linkage disequilibrium (LD) as measured by the pairwise r^2 value between adjacent markers was higher in the biparental populations than in the mixed population and increased with larger values of N_M (Table 1). The LD was greater than 0.70 for all biparental populations. In the mixed barley population, LD at the highest marker density was 0.53.

As expected from Eq. 1, r_{MP} increased as N increased (Suppl. Tables 1–5). For example, in the biparental maize population and with the highest N_M (1213 markers) and $h^2 = 0.30$, the prediction accuracy for grain yield was $r_{MP} = 0.19$ with $N = 48$, $r_{MP} = 0.26$ with $N = 96$, and $r_{MP} = 0.33$ with $N = 192$. In the mixed wheat population and with the highest N_M (731 markers) and $h^2 = 0.30$, the prediction accuracy for heading date was $r_{MP} = 0.40$ with $N = 48$, $r_{MP} = 0.43$ with $N = 72$, and $r_{MP} = 0.46$ with $N = 96$.

Similar findings regarding the effects of N_M and N on r_{MP} were obtained in previous studies. In biparental populations of maize, Arabidopsis, barley, and wheat, the

highest N_M generally resulted in the highest accuracy and the highest N always resulted in the highest accuracy (Lorenzana and Bernardo 2009; Guo et al. 2011; Heffner et al. 2011b). Similarly, mixed populations in wheat (Heffner et al., 2011a), forest trees (Grattapaglia and Resende, 2011), and maize (Albrecht et al., 2011) showed that increasing N and N_M increased prediction accuracy.

Influence of Heritability

Traits with high unmodified h^2 (for biparental populations) or g^2 (for mixed populations) generally had high r_{MP} relative to other traits in that population (Table 2, Suppl. Tables 1–5). There were a few exceptions to this trend; for example, in the maize biparental population, root lodging had the highest r_{MP} but also had the second lowest h^2. While Eq. 1 suggests that a higher h^2 should always lead to higher r_{MP}, my findings are consistent with previous research that shows most traits with high h^2 are predicted well, but that there are exceptions (Grattapaglia et al., 2009; Heffner et al., 2011a; Heffner et al. 2011b; Albrecht et al., 2011). For example, in the wheat biparental population Cayuga × Caledonia, grain softness had an h^2 of 0.88 and prediction accuracy of 0.37 whereas sucrose solvent retention had a much lower h^2 of 0.45 but a prediction accuracy of 0.41 (Heffner et al., 2011b).

Within a given trait, reducing the h^2 or g^2 almost always resulted in reductions in r_{MP} (Fig. 1, Suppl. Tables 1–5). There was one trait in the wheat mixed population, heading date, that showed a significant increase in r_{MP} at the highest N_M and N when h^2 was decreased from the original value of $h^2 = 0.95$ ($r_{MP} = 0.45$) to 0.50 ($r_{MP} = 0.49$) (Fig. 1). There is no clear explanation for this finding. The steepness of the decrease in r_{MP} as h^2 or g^2 decreased also differed among traits. For example, in the barley mixed

population, reduction in the g^2 of grain protein resulted in a steep decline in r_{MP}, whereas decreasing the g^2 of plant height or heading date resulted in relatively little change in r_{MP}.

The estimates of h^2 were obviously subject to sampling variation. For example, the estimates of h^2 and their 90% confidence intervals (in parentheses) in the maize biparental population were h^2 = 0.45 (0.33, 0.54) for root lodging and h^2 = 0.44 (0.33, 0.53) for grain yield (Table 2). I took the estimates of h^2 and added nongenetic effects with a variance of V_{Extra} to reduce the h^2 to 0.30 and 0.20. Now suppose the true values were h^2 = 0.33 (i.e., lower limit of confidence interval) for root lodging and h^2 = 0.53 (i.e., upper limit of confidence interval) for grain yield. In this situation, the target h^2 of 0.30 would have corresponded to an actual h^2 of 0.22 for root lodging and 0.36 for grain yield. Some caution is therefore needed in interpreting the results. On the other hand, most of the traits had h^2 estimates that were well outside each other's confidence intervals. For example, lodging in the barley biparental population had h^2 = 0.67 (0.59, 0.73), and it was extremely unlikely that the true value of h^2 for lodging was equal to that of alpha amylase [h^2 = 0.82 (0.86, 0.88)] or malt extract [h^2 = 0.88 (0.86, 0.90)].

Importance of Trait

Equation 1 indicates that the product of h^2 and N, rather than h^2 and N individually, is the key factor that determines prediction accuracy. I found that for the same trait within a population, r_{MP} values generally were not different when Nh^2 was constant. For example, in the biparental maize population, the r_{MP} for moisture was 0.30 with both N = 72 and h^2 = 0.50, and N = 180 and h^2 = 0.20 (Nh^2 = 36). Similarly, in the mixed wheat population, the r_{MP} for maturity was not significantly different with N = 72 and g^2 = 0.50 (r_{MP} = 0.41) and with N = 120 and g^2 = 0.20 (r_{MP} = 0.42; Ng^2 = 36). There

were three instances (simulated population with 10 QTL and 50 QTL, and lodging in the barley biparental population) where r_{MP} differed significantly for different combinations of N and h^2 that led to the same Nh^2. In these three instances, the differences in r_{MP} were only 0.02–0.03. These results support the validity of Eq. 1 and indicate that, for the same trait within the same population, a decrease in h^2 can be compensated by a proportional increase in N (and vice-versa) so that r_{MP} is maintained.

In contrast, across different traits within the same population, holding N, h^2 (or g^2), and N_M constant did not lead to the same r_{MP}. In the maize biparental population, r_{MP} was consistently lower for grain yield than for the other traits even when N, h^2, and N_M were constant across traits (Fig. 1). Likewise, grain yield in the barley biparental population and grain yield and biomass yield in the wheat mixed population had lower r_{MP} compared with the other traits. Across populations, most of the traits studied could be grouped into four categories: yield (both grain and biomass), flowering time, height, and lodging. The results indicated that just as h^2 tends to be lowest for yield, r_{MP} is also lowest for yield traits even when its h^2 is as high as that for other traits. Plant height and lodging were always predicted most accurately, followed by flowering time (Table 2, Suppl. Tables 1–5).

In addition to N and h^2 (and assuming that N_M is large so that the genome is saturated with markers), the additional factor affecting the expected prediction accuracy in Eq. 1 is M_e, the effective number of chromosome segments (Daetwyler et al., 2008; 2010). Assuming the genome comprises k chromosomes that each are L Morgans in length, M_e has been proposed as equal to $2N_eLk/\log(N_eL)$ (Goddard and Hayes, 2011), where N_e is the effective population size. The N_e for the biparental populations was 1, i.e.,

the recombinant inbreds were all descended from a single non-inbred plant (i.e., the F_1). The use of $N_e = 1$ in the above equation for M_e fails to give a positive M_e. As an alternative, I considered M_e as equal to the number of chromosomes (low M_e), the size of the linkage map divided by 50 cM (medium M_e), and N_M (high M_e). I then used these M_e values in Eq. 1 and multiplied the result by h to obtain the predicted r_{MP} (Table 2). In nine instances out of the 22 population-trait combinations, the observed r_{MP} fell between the predicted r_{MP} for the low M_e and the predicted r_{MP} for the medium M_e. In 12 instances, the observed r_{MP} fell between the predicted r_{MP} for the medium M_e and the predicted r_{MP} for the high M_e. Traits in the mixed populations tended to have an r_{MP} between the predicted r_{MP} values for the medium and high M_e, and this result was consistent with an increase in the number of independent chromosome segments as LD decreases. Grain yield in the mixed wheat population had r_{MP} below any of the predicted r_{MP}. The differences in r_{MP} despite N, h^2, and N_M being held constant lead us to speculate that M_e must not simply be a function of N_e and the size of the genome (Goddard and Hayes, 2011), but it must also be a function of the number of QTL. In this study, a trait controlled by 50 QTL was predicted the most accurately, followed by a trait controlled by 10 QTL, and lastly a trait controlled by 100 QTL (Suppl. Table 5). However, the differences in r_{MP} with varying numbers of QTL were much smaller than the differences in r_{MP} for different traits in the empirical populations. The lower r_{MP} with 10 QTL than with 50 QTL may be due to the RR-BLUP approach not being optimal when only a few QTL control the trait (Meuwissen et al., 2001; Lorenz et al., 2011; Resende et al., 2012). Previous research showed that in a barley mixed population, a simulated trait controlled by 20 QTL was generally predicted with greater accuracy than one controlled by 80 QTL

(Zhong et al., 2009). In forest trees, accuracy of genomewide selection declined as more QTL controlled the trait (Grattapaglia et al., 2009).

Implications

In practice, breeders typically select for multiple traits that differ in their genetic architecture and h^2. If the same training population is used for all traits, breeders must then be prepared to accept that r_{MP} will be lower for some traits than for other traits, in the same way that h^2 is lower for some traits than for others. On the other hand, traits with initially low h^2 can be evaluated with larger N or the h^2 for a subset of traits can be increased by the use of additional testing resources. This practice is illustrated in the barley biparental population: extract and alpha amylase, which have high h^2 but are expensive to measure, were evaluated at nine locations whereas grain yield, which has low h^2 but is simpler to measure, was evaluated at 16 environments (Hayes et al., 1993).

While there has been much research on the influence of genetic architecture on QTL mapping (Holland, 2007) and association mapping (Myles et al., 2009), further studies are needed on why some traits are predicted more accurately than others in genomewide prediction (Meuwissen, 2012). In particular, further studies are needed to determine M_e. Also, while epistasis may be involved, previous results for the same maize and barley datasets showed that attempting to account for epistasis did not lead to better predictions (Lorenzana and Bernardo, 2009). Due to the importance of the trait on prediction accuracy, accumulated empirical data on the r_{MP} for different traits will be crucial to the successful design of training populations for genomewide selection.

Table 1 Number of single nucleotide polymorphism markers, spacing between adjacent markers, and linkage disequilibrium (r^2) for the low, medium, and high density marker sets in each population.

Population	Size of linkage map	High density			Medium density			Low density		
		N_M †	Spacing ‡	r^2 §	N_M	Spacing	r^2	N_M	Spacing	r^2
	cM		cM			cM			cM	
Maize biparental population	6240	1213	5	0.72	512	12	0.55	256	24	0.37
Barley biparental population	1250	223	6	0.80	100	13	0.63	48	26	0.27
Barley mixed population	1250	1178	1	0.53	768	2	0.48	384	3	0.44
Wheat mixed population	2569	731	4	.	576	4	.	384	7	. ¶
Simulated population	1749	350	5	0.82	140	12	0.61	70	25	0.36

† Number of single nucleotide polymorphism markers used

‡ Approximate spacing (in centiMorgans, cM) between adjacent markers

§ Linkage disequilibrium as estimated by the mean pairwise r^2 values between adjacent markers. Linkage disequilibrium could not be estimated in the wheat mixed population.

Table 2 Heritability (h^2), observed genomewide prediction accuracy (r_{MP}), and predicted r_{MP} assuming different effective numbers of chromosome segments (M_e) for different traits in different populations. Low M_e was equal to the number of chromosomes, medium M_e was equal to the genome size divided by 50, and high M_e was equal to the number of markers.

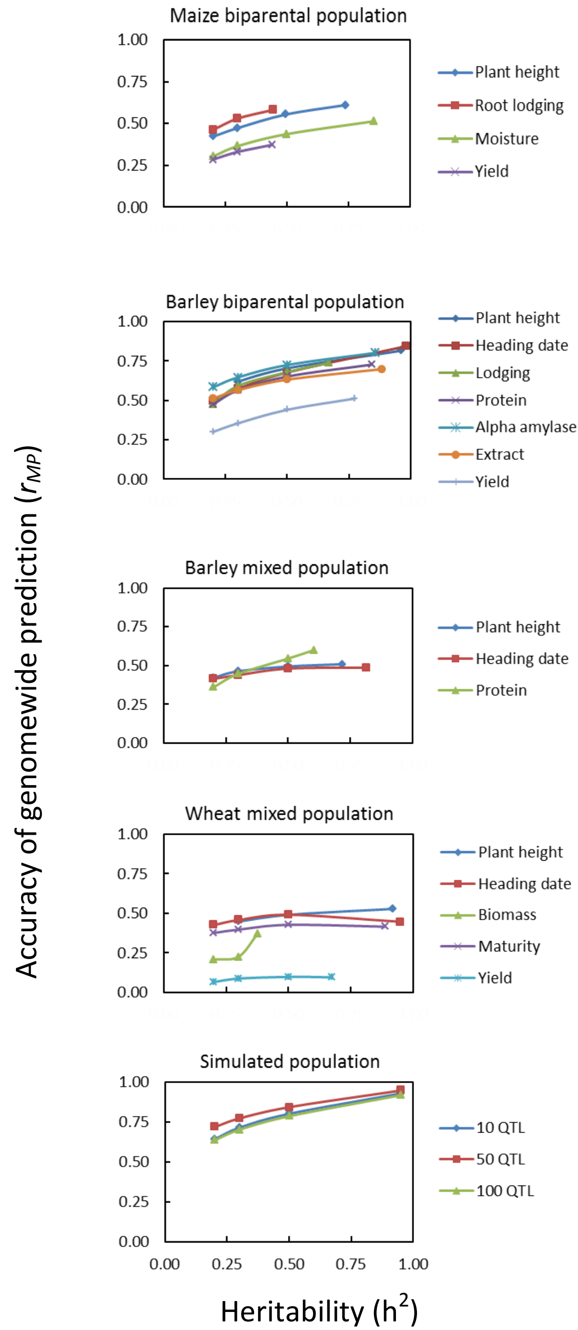
Population and trait	h^2	CI †	r_{MP} ‡	Predicted r_{MP}		
				Low M_e	Medium M_e	High M_e
Maize biparental population						
Plant height	0.74	(0.69, 0.78)	0.61	0.83	0.52	0.28
Root lodging	0.45	(0.33, 0.54)	0.58	0.63	0.34	0.17
Moisture	0.85	(0.82, 0.88)	0.51	0.90	0.58	0.32
Yield	0.44	(0.33, 0.53)	0.37	0.63	0.33	0.17
Barley biparental population						
Plant height	0.96	(0.95, 0.97)	0.82	0.94	0.84	0.53
Heading date	0.98	(0.98, 0.98)	0.84	0.96	0.85	0.54
Lodging	0.67	(0.59, 0.73)	0.74	0.78	0.66	0.39
Protein	0.84	(0.81, 0.87)	0.73	0.88	0.77	0.47
Alpha amylase	0.86	(0.82, 0.88)	0.80	0.89	0.78	0.48
Extract	0.88	(0.86, 0.90)	0.70	0.90	0.80	0.49
Yield	0.77	(0.72, 0.81)	0.51	0.84	0.73	0.44
Barley mixed population						
Plant height	0.72	(0.61, 0.80)	0.51	0.81	0.70	0.20
Heading date	0.82	(0.74, 0.87)	0.49	0.87	0.76	0.23
Protein	0.61	(0.45, 0.72)	0.60	0.74	0.62	0.17

Wheat mixed population						
Plant height	0.92	(0.90, 0.94)	0.53	0.86	0.72	0.32
Heading date	0.95	(0.94, 0.96)	0.45	0.88	0.74	0.32
Maturity	0.89	(0.86, 0.91)	0.42	0.84	0.70	0.30
Biomass	0.38	(0.22, 0.51)	0.37	0.49	0.36	0.13
Yield	0.68	(0.60, 0.75)	0.10	0.72	0.58	0.24
Simulated population						
10 QTL	0.95		0.93	0.95	0.83	0.57
50 QTL	0.95	§	0.95	0.95	0.83	0.57
100 QTL	0.95		0.92	0.95	0.83	0.57

† 90% confidence interval on estimates of h^2. True values of h^2 were known in the simulated population.

‡ From cross validation with the largest training population size (N) and number of markers (N_M) in each population.

Figure 1: Accuracy of genomewide prediction (r_{MP}) with different levels of heritabilty (h^2). Results are for the highest marker density and training population size within each population.



Chapter 2: Marker imputation prior to genomewide prediction among diverse maize inbreds and elite barley inbreds

In genomewide prediction, marker imputation reduces costs by genotyping the training population with M_{Total} single nucleotide polymorphism (SNP) markers, genotyping the test population with M_{Low} markers, and predicting the genotypes at the $M_{Total} - M_{Low}$ markers in the test population. My objectives were to determine (i) how the type of inbred collection affects imputation accuracy; (ii) if genomewide prediction with imputed markers can be as accurate as genomewide prediction with non-imputed markers; and (iii) if lower values of M_{Total} can increase imputation accuracy without decreasing genomewide prediction accuracy (r_{MP}). Imputation was less accurate in a diverse collection of 283 historical maize (*Zea mays* L.) inbreds, genotyped with 30,973 markers, than among 860 elite barley (*Hordeum vulgare* L.) inbreds genotyped with 2019 SNPs. Without imputation, r_{MP} plateaued between 288 and 6000 SNPs in the maize collection and between 96 and 2019 SNPs in the barley collection. With imputation, many combinations of M_{Total} and M_{Low} led to r_{MP} values that were as high as the r_{MP} with M_{Total} non-imputed markers. Because smaller values of M_{Total} often had high imputation accuracy and did not lead to significantly lower r_{MP}, imputing to the highest number of markers is not generally recommended in inbred collections.

Introduction

Genomewide selection (or genomic selection) allows marker-assisted selection without quantitative trait locus (QTL) mapping (Meuwissen et al., 2001). In genomewide selection, marker effects at a large number of markers are estimated from a training population which has been genotyped and phenotyped. The marker effects are then used to predict the performance of individuals in a test population which has been genotyped but not phenotyped. In plants, a training population could comprise a single biparental cross (Lorenzana et al., 2009; Heffner et al., 2011b), multiple biparental crosses (Albrecht et al., 2011; Massman et al., 2012), or a collection of diverse inbreds or accessions (Meuwissen et al., 2001; Heffner et al., 2011a; Grattapaglia et al., 2011). Inbred collections tend to have weaker linkage disequilibrium (LD) between the markers and QTL so large numbers of markers are subsequently needed for genomewide prediction when an inbred collection is used as a training population (Heffner et al., 2011a; Lorenz et al., 2011).

Genotyping can be prohibitively expensive when many markers are needed to obtain accurate genomewide predictions (Boichard et al., 2012). Marker imputation reduces costs by genotyping the training population with many markers (M_{Total}), genotyping the test population with fewer markers (M_{Low}), and predicting the genotypes at the $M_{Total} - M_{Low}$ markers in the test population by an imputation process. Software such as fastPHASE (Scheet et al., 2006), BEAGLE (Browning, 2007), IMPUTE (Marchini et al., 2007), and PLINK (Purcell et al., 2007) use localized patterns of LD among the M_{Total} markers to identify haplotype blocks (Nothnagel et al., 2009). Based on

the common markers between the M_{Total} and M_{Low} sets of markers, the haplotype blocks are used to impute the $M_{Total} - M_{Low}$ markers. Imputation can also integrate pedigree information to further improve accuracy (Hayes et al., 2011).

Marker imputation has been found accurate in sheep (*Ovis aries* L.; Hayes et al., 2011) and in dairy and beef cattle (*Bos taurus* L.; Dassonneville et al., 2011; Wang et al., 2012; Weigel et al., 2010). Such accurate imputations have led to negligible decreases in the accuracy of subsequent genomewide prediction in animals (Berry et al., 2011; Dassonneville et al., 2011). In a cattle population with a low density panel of 2730 single nucleotide polymorphism (SNP) markers and a high density panel of 51,602 SNPs, imputation from 2730 to 51,602 SNPs led to a negligible difference in genomewide prediction accuracy compared with having all the cattle genotyped with 51,602 SNPs (Berry et al., 2011). Marker imputation prior to genomewide prediction is therefore becoming common in livestock breeding programs where it is cost-prohibitive to genotype all individuals with high-density SNP chips (Cleveland et al., 2011).

In plants, marker imputation has been found effective in barley (*Hordeum vulgare* L.) and maize (*Zea mays* L.) where the accuracy of imputation was affected by the level of LD in the population, minor allele frequency of the markers, and genetic relatedness within the population (Jannink et al., 2009; Hickey et al., 2012). Imputation accuracy is commonly measured as the correlation between the true and imputed marker genotypes (Hickey et al., 2012). The level of accuracy that is acceptable would depend on how the imputed marker information will be used, but accuracy greater than 0.80 to 0.90 is generally considered sufficient (Jannink et al., 2009). In barley, imputation with M_{Total} =

3000 SNPs and M_{Low} = 600 SNPs led to 93% of the (3000 – 600) = 2400 markers being imputed with an accuracy greater than 0.89 (Jannink et al., 2009). In maize, imputation with M_{Total} = 35,081 SNPs and M_{Low} = 8,774 SNPs led to a mean imputation accuracy of 0.87 (Hickey et al., 2012).

These previous studies in plants have focused on using different numbers of M_{Low} but the effect of different numbers of M_{Total} markers on imputation accuracy is yet to be explored. Furthermore, the effect of marker imputation on the accuracy of subsequent genomewide predictions has not been tested in plants. Suppose imputing from 384 to 55,000 SNPs is found ineffective but imputing from 384 to 1536 SNPs is effective. In this situation, marker imputation is useful if genomewide predictions are equally accurate with 1536 and 55,000 SNP markers but is not useful if more than 1536 SNPs are needed to obtain accurate genomewide predictions. Lastly, differences in the composition and population structure of inbred collections (e.g., diverse panel versus elite breeding germplasm) lead to differences in the level of LD and number of markers needed for full genome coverage. Information will be useful on how the type of inbred collection used as the training population affects the accuracies of marker imputation and subsequent genomewide prediction.

My objectives were to determine (i) how the accuracy of imputation is affected by the type of inbred collection used as the training population; (ii) if genomewide prediction with imputed markers can be as accurate as genomewide prediction with non-imputed markers in inbred collections; and (iii) if lower values of M_{Total} can be used to increase imputation accuracy without decreasing genomewide prediction accuracy.

Materials and Methods

Overview

I considered a diverse maize inbred collection (Schaefer et al., 2012) and a contemporary barley inbred collection (http://hordeumtoolbox.org). For each species, I compared two types of training and test populations in which M_{Total} varied: non-imputed population in which marker imputation was not performed, and imputed population in which marker imputation was performed. When imputation was performed, the training population had M_{Total} markers, the test population had M_{Low} markers, and the program fastPHASE (Scheet et al., 2006) was used to impute the $M_{Total} - M_{Low}$ marker genotypes in the test population so that the test population had data for all M_{Total} markers (Fig. 2). For the imputed populations, M_{Low} ranged from 96 to 6000 SNPs in maize and from 96 to 1536 SNPs in barley whereas M_{Total} ranged from 1536 to 30,973 SNPs in maize (Table 3) and from 768 to 2019 SNPs in barley (Table 4). Genomewide predictions were subsequently obtained by ridge-regression best linear unbiased prediction (RR-BLUP; Meuwissen et al., 2001), and the accuracy of genomewide predictions was assessed by calculating the correlation between observed and predicted performance for each trait.

Genotypic and phenotypic data

The maize collection of 283 inbreds included 143 historical inbreds developed at the University of Minnesota and 140 publicly or privately developed inbreds with B73, Mo17, A158/Oh43, A321, and PH207 genetic backgrounds (Schaefer et al., 2012). The 283 inbreds were evaluated at five Minnesota locations in 2011 for kernel composition (protein, oil, and starch), plant and ear height, and days to silking and anthesis. Best

linear unbiased estimates of the mean performance of each inbred for each trait were calculated using PROC MIXED in the SAS software version 9.2 for Windows 7 (SAS Institute; Cary, NC) to account for variation due to incomplete blocks and replications. For SNP analysis, leaf tissue samples were collected in summer 2011 at the St. Paul, MN location and sent to DNA LandMarks (Saint-Jean-sur-Richelieu, Quebec) for analysis with the Maize SNP50 Beadchip (Illumina, San Diego, CA), which included 56,110 SNP markers.

The barley collection of 860 inbreds included advanced inbreds from nine breeding programs (Busch Agriculture, University of Idaho, Montana State University, North Dakota State University two-row, North Dakota State University six-row, University of Minnesota, Utah State University, Virginia Tech, and Washington State University) with 95-96 inbreds per breeding program. The barley inbreds were evaluated in Pullman, WA in 2009 for grain hardness, grain width, grain weight, polyphenol oxidase activity and hull proportion. Genotypic data for 3072 SNPs were downloaded from the Hordeum Toolbox (available at http://hordeumtoolbox.org, accessed April 14, 2012). Hull proportion was measured in 762 of the inbreds as hull-less inbreds were not evaluated for this trait.

SNPs with more than 10% missing data, a minor allele frequency less than or equal to 7%, or unknown chromosome or map position were removed from both datasets, leaving 30,973 SNPs for maize and 2019 SNPs for barley. Marker genotypes were coded as homozygous major allele (1), homozygous minor allele (–1), and heterozygous (0).

Marker-based coefficients of coancestry were estimated in the maize and barley collections based on the pairwise simple matching coefficient (Bernardo, 1993). I assumed that the two inbreds with the minimum similarity in each collection were unrelated and used this minimum similarity as an estimate of the probability that two inbreds have marker alleles that are alike in state but not identical by descent (Lynch, 1988).

Selection of M_{Low} marker sets

I evaluated subsets of M_{Low} = 96, 192, 288, 384, 768, 1536, 3000, 6000, 9000, 12,000 and 14,850 SNPs in maize and M_{Low} = 96, 192, 288, 384, 768, and 1536 SNPs in barley (Fig. 2). The M_{Low} subsets in maize were chosen based on their physical position (Ganal et al., 2011; Chia et al., 2012) and the M_{Low} subsets in barley were chosen based on their linkage map position (Close et al., 2009). The markers were then chosen to balance even marker spacing and high minor allele frequency (Zhang et al., 2010). The number of markers retained on each chromosome was calculated as M_{Low} multiplied by the size of the chromosome and divided by the size of the genome. Next, each chromosome was divided into bins by dividing the physical or map size of the chromosome by the number of markers for that chromosome. The M_{Low} markers then included one marker from each bin. In the first bin, the marker with the highest minor allele frequency was chosen. All other bins had the markers ranked according to $S_i = M_i (z - |z - d_i|)$, where S_i was the score of the ith marker, M_i was the minor allele frequency, z was the size of the bin, and d_i was the difference between the position of the last chosen marker and the ith marker. After scoring each marker in a bin, the marker with the

highest S_i was chosen for the M_{Low} marker subset. At high densities, sometimes no markers were in a bin. In that situation, the bin size was increased by z until markers were available. Sets of M_{Low} markers were chosen independently for each level of marker coverage; for example, not every marker in the 192 SNP set was necessarily in the 288 SNP set.

For imputation, the maize and barley inbred collections were each partitioned so that a random set of inbreds had genotypic data for the M_{Total} markers and the remaining inbreds had data for M_{Low} markers, with the remaining $M_{Total} - M_{Low}$ markers being assumed unknown. The same sets of random inbreds were assigned to retain M_{Total} or M_{Low} for all M_{Total} and M_{Low} combinations. One-third of each collection (95 maize inbreds and 288 barley inbreds) had M_{Total} markers whereas the remaining two-thirds of the collection (188 maize inbreds and 572 barley inbreds) had M_{Low} markers.

Linkage disequilibrium

The LD was measured (in the non-imputed populations only) as the mean r^2 between adjacent markers across the genome. The LD in the entire maize inbred collection and the entire barley inbred collection was calculated with Haploview version 4.2 (Barrett et al., 2005). For a given number of M_{Total} markers, the LD between adjacent markers was calculated for each chromosome and averaged across all chromosomes. Markers with the same physical or map position were excluded. In the maize collection, two markers were excluded at M_{Total} = 6000 and 14 markers were excluded at M_{Total} = 30,973. Because only linkage map positions were available in the barley collection, many more markers were excluded leaving the following number of markers with unique

positions at each value of M_{Total} (M_{Total} in parentheses): 96 (96), 189 (192), 278 (288), 367 (384), 539 (768), 705 (1536) and 768 (2019).

Marker imputation

The program fastPHASE version 1.4.0 (Scheet et al., 2006), run on the University of Minnesota Supercomputing Institute Linux server, was used for all imputations in this study. The fastPHASE software identifies clusters of shared haplotypes along the chromosome using a hidden Markov model to identify blocks of LD among inbreds with M_{Total} markers, and these blocks are then used to impute the missing genotypes in inbreds with M_{Low} markers (Scheet et al., 2006). I chose fastPHASE because it is commonly used (Jannink et al., 2009; Iwata et al., 2010; Weigel et al., 2010) and does not require pedigree information. The fastPHASE software was run with 10 starts of the EM algorithm and haplotypes were not phased. Because pedigree information was not available for either dataset, imputation was based only on LD. Imputation was performed 100 times per imputation scheme and the results were averaged (Iwata et al., 2010). Therefore, not all genotypes were –1, 0, or 1 and non-whole numbers were possible as coded genotypes. Computation time varied depending on the inbred collection and M_{Total}; the value of M_{Low} did not noticeably change computation time. On the low end, a single run of fastPHASE for one chromosome took about 10 minutes per chromosome in the barley collection with M_{Total} = 768 or in the maize collection with M_{Total} = 1536. On the high end, one run of fastPHASE for one chromosome took about 140 minutes in the maize collection with M_{Total} = 30,973.

Imputation accuracy (r_I) was calculated as the correlation coefficient between the mean imputed and true marker genotypes among inbreds for which imputation was performed. Imputation accuracy is reported as the mean r_I across inbreds.

Genomewide prediction accuracy

Two types training and test populations were used for genomewide predictions: (i) non-imputed training and test populations with M_{Total} = 96–2019 SNPs in the barley collection (Table 4) and M_{Total} = 96–30,973 SNPs in the maize collection (Table 3); and (ii) non-imputed training population with M_{Total} markers and imputed test population with M_{Low} markers. In the imputed test population, genotypes at the $M_{Total} - M_{Low}$ markers were the means from 100 runs of fastPHASE (Iwata et al., 2010). Across the different values of M_{Total} and M_{Low}, the same set of random inbreds was used as the training population with the remainder being the test population.

Marker effects were predicted by RR-BLUP as implemented in the R package rrBLUP version 3.8 (Endelman, 2011). Marker effects were multiplied by the coded genotypes (ranging from –1.0 to 1.0) of the test population and summed for each inbred to obtain predicted performance. With imputation, the marker design matrix included non-integers for the $M_{Total} - M_{Low}$ imputed genotypes. The accuracy of prediction (r_{MP}) was the correlation between predicted values and observed phenotypic values. The significance of differences in r_{MP} was tested based on Fisher's z transformation ($P = 0.05$) (Fisher, 1915).

To assess what would happen if different sets of inbreds were randomly assigned to the training and test populations, I chose combinations of M_{Low} and M_{Total} for which

most traits had r_{MP} not significantly different from the highest observed r_{MP}. Four additional random assignments of inbreds to the training and test populations were then considered for a total of five repeats (Tables 5 and 6). As before, each training population had M_{Total} markers and the test population had M_{Low} markers. In the imputed test population, genotypes at the $M_{Total} - M_{Low}$ markers were the means from 100 runs of fastPHASE. Subsequent procedures for imputation and genomewide prediction were as previously described. Mean values for r_I and r_{MP} across the five repeats were calculated for each combination of M_{Low} and M_{Total}. The sampling variance, which was then used to obtain least significant differences (LSD; P = 0.05), was calculated with SAS PROC GLM by treating r_{MP} as the dependent variable and M_{Total}, M_{Low}, and repeats as the independent variables.

Heritability is undefined in a collection of inbreds that are not members of the same random-mating population. I define Θ as $\Sigma\tau_i^2/(N - 1)$, where τ_i is the effect of the *i*th inbred and *N* is the number of inbreds. The ratio between Θ and the total phenotypic variance indicates how much of the observed variation is due to genetic causes. I calculated this ratio, which I refer to as g^2, for each trait using a mixed model for the maize collection where inbreds had fixed effects and replication and block had random effects. For the barley collection a general linear model was used where inbreds had a fixed effect. All calculations were done using SAS software.

Results and Discussion

Linkage disequilibrium and imputation accuracy

In the maize collection, LD, which was measured as the mean r^2 between adjacent markers, increased steadily from 0.03 with 96 SNPs to 0.32 with all 30,973 SNPs (Table 3). In the barley collection, LD did not increase much beyond M_{Total} = 288 SNPs (LD = 0.31), with LD values of 0.30 with 384 SNPs, 0.31 with 768 SNPs, 0.28 with 1536 SNPs, and 0.24 with 2019 SNPs (Table 4). Because many markers needed to be excluded in the barley collection due to their identical linkage map positions, the estimated LD at higher values of M_{Total} in the barley collection may be misleading. Nevertheless, LD in the maize collection at the highest marker density (M_{Total} = 30,973) was similar to LD in the barley collection at the higher marker densities ($M_{Total} \geq 288$) (Tables 1 and 2).

Imputation accuracy (r_I), which was the correlation between imputed marker genotypes and observed marker genotypes averaged across inbreds for which imputation was performed, was poor when imputing from low levels of M_{Low} and was generally lower in the maize collection than in the barley collection. In addition, r_I decreased as the number of markers to be imputed (i.e., $M_{Total} - M_{Low}$) increased. Imputation accuracy ranged from 0.04 (M_{Low} = 96, M_{Total} = 30,973) to 0.75 (M_{Low} = 6000, M_{Total} = 30,973) in the maize collection (Table 3) and from 0.52 (M_{Low} = 96, M_{Total} = 2019) to 0.96 (M_{Low} = 1536, M_{Total} = 2019) in the barley collection (Table 4). When LD was similar between the two collections, r_I was higher in the barley collection than the maize collection suggesting that the mean r^2 between adjacent markers was not a reliable predictor of r_I.

There was no combination of M_{Total} and M_{Low} for which r_I was higher in maize than in barley.

The difference in r_I between the maize and barley collections is explained by differences in the levels of LD, relatedness among inbreds, and numbers of inbreds with M_{Total} and M_{Low} markers. At lower levels of M_{Total}, LD was higher in the barley collection than in the maize collection. Because imputation works by identifying localized patterns of LD, stronger LD leads to more accurate imputation (Iwata et al. 2010). The relatedness of inbreds or accessions in a population and the number of close relatives within a population is also known to increase r_I (Zhang et al., 2010; Hickey et al., 2012). The higher mean coefficient of coancestry, estimated using marker data, in the barley collection (0.36) than in the maize collection (0.28) indicated that the barley inbreds were more closely related than the maize inbreds. While the barley inbreds were elite breeding lines, the maize inbreds included diverse lines from the defunct University of Minnesota maize inbred development program, multiple public breeding programs, and multiple breeding companies. Furthermore, the barley collection was much larger than the maize collection and r_I is known to increase as more inbreds are genotyped at a high marker density (Hayes et al., 2011).

Accuracy of genomewide prediction with and without imputation

Without imputation, the accuracy of genomewide prediction (r_{MP}) plateaued between 288 and 6000 SNPs in the maize collection (Table 3) and between 96 and 2019 SNPs in the barley collection (Table 4) depending on the trait. In the maize collection, the r_{MP} with 384 or more SNPs was not significantly lower than the highest r_{MP} observed for

all traits. Except for one trait (polyphenol oxidase activity), the r_{MP} with 768 or more SNPs in the barley collection was not significantly lower than the highest r_{MP} observed.

Imputed and non-imputed populations frequently had r_{MP} values that did not differ significantly. In the maize collection, imputing from $M_{Low} \geq 3000$ to M_{Total} = 30,973 led to r_{MP} values that were not significantly lower than the highest r_{MP} attained. In the barley collection, imputing from $M_{Low} \geq 768$ to $M_{Total} = 2019$ led to r_{MP} values that were not significantly lower than the highest r_{MP} observed. The r_{MP} values with imputed SNPs were not significantly lower than the highest r_{MP} when r_I exceeded 0.51 in the maize collection and exceeded 0.90 in the barley collection.

For a given M_{Total}, an increase in M_{Low} led to higher r_{MP}. For example, with M_{Total} = 1536 in the maize collection, r_{MP} for anthesis date increased from 0.31 ($M_{Low} = 96$) to 0.62 ($M_{Low} = 288$). With $M_{Total} = 1536$ in the barley collection, r_{MP} for grain hardness increased from 0.44 ($M_{Low} = 96$) to 0.63 ($M_{Low} = 288$). For a given M_{Low}, a decrease in M_{Total} generally resulted in an increase in r_{MP}. For example, with $M_{Low} = 192$ in the maize collection, r_{MP} for ear height decreased from 0.35 ($M_{Total} = 1536$) to 0.32 ($M_{Total} = 3000$), 0.32 ($M_{Total} = 6000$), and 0.26 (M_{Total} = 30,973). With $M_{Low} = 96$ in the barley collection, r_{MP} for grain width was 0.81 with $M_{Total} = 768$, 0.79 with $M_{Total} = 1536$, and 0.78 with $M_{Total} = 2019$.

Due to the required computing time, the results in Tables 3 and 4 were for a single repeat, i.e., one partitioning of the inbreds into the training population and test population. The results of four additional repeats supported the results of the first repeat. The mean r_I values across five repeats in the maize collection (Table 5) and barley

collection (Table 6) did not differ significantly from the r_I for a single repeat (Tables 3 and 4).

Across repeats and for most traits, imputation led to r_{MP} values that were significantly higher than the r_{MP} in non-imputed populations with low M_{Total} (Tables 3 and 4). For most traits, imputation prior to genomewide prediction did not significantly decrease r_{MP} compared to using high values of M_{Total} in non-imputed populations. Overall, the highest values of r_{MP} were obtained with M_{Low} = 288 and M_{Total} = 3000 (mean r_I = 0.39) in the maize collection (Table 5), and with M_{Low} = 288 and M_{Total} =1536 (mean r_I = 0.86) in the barley collection (Table 6). The combinations of M_{Total} and M_{Low} that led to the highest r_{MP} did not have the highest r_I; in fact, the combinations of M_{Total} and M_{Low} with the highest r_{MP} had the lowest r_I across repeats for both the barley and maize collections (Tables 3 and 4).

Use of marker imputation in breeding programs

Overall, I found that while genomewide prediction accuracy (r_{MP}) can be equal with or without imputation, r_{MP} with imputation is affected by the type of inbred collection used, accuracy of imputation (r_I), total number of markers used, and trait being predicted. The r_I was higher in the elite barley collection than in the diverse maize collection. While higher r_I generally led to higher r_{MP}, this result was not always observed. Lower values of M_{Total} increased r_I and did not always result in lower r_{MP}. These results indicate that imputation prior to genomewide prediction would be useful in plant breeding programs.

A key finding from this study is the importance of imputing to the number of M_{Total} markers where r_{MP} begins to plateau, rather than to the full number of available markers. Selecting the minimal M_{Total} improves imputation accuracy and decreases costs. Some breeding programs are currently considering switching from a chip genotyping system to genotyping by sequencing (Poland et al., 2012). I believe that chip genotyping is still a valuable strategy because relatively few markers (few thousands instead of tens of thousands) are generally needed to obtain accurate genomewide predictions in plants (Lorenz et al., 2011). Use of chip genotyping also avoids some of the pitfalls of genotyping by sequencing. In contrast to genotyping by sequencing, chip genotyping has lower costs, can capture heterozygosity, requires less data handling, and produce genotypes with greater accuracy (Davey et al., 2011; Neilson et al., 2011).

Breeders select for multiple traits, and the effectiveness of imputation prior to genomewide prediction depended on the trait. For example, for polyphenol oxidase activity in the barley collection, only three combinations of M_{Total} and M_{Low} in imputed populations led to r_{MP} values that were not significantly lower than the highest r_{MP} in non-imputed populations (Table 4). In contrast, for grain width, 18 of the 20 combinations of M_{Total} and M_{Low} did not lead to a significant decrease in r_{MP}. The results for polyphenol oxidase activity were the exception rather than the norm, and the underlying reasons for this result are unknown.

I have assumed that, in practice, the training population is genotyped with M_{Total} markers, the test population is genotyped with M_{Low} markers, and the genotypes at the $M_{Total} - M_{Low}$ markers in the test population are to be imputed. In general, imputation

prior to genomewide prediction is most effective when r_I is high and lower values of M_{Total} are used. This study did not provide clear guidelines on what level of r_I is sufficient: an r_I of at least 0.80 seemed sufficient in the barley collection and an r_I of at least 0.40 seemed sufficient in the maize collection. Instead of having a threshold r_I above which imputation prior to genomewide prediction is deemed useful, an alternative approach is to use a cross-validation scheme on an ad hoc basis. If phenotypic data and marker data on M_{Total} SNPs are available for the training population, cross-validation analysis similar to that described in this study could be conducted to determine the minimum M_{Total} needed for accurate genomewide predictions within the training population and the effect on r_{MP} of imputing from M_{Low} to M_{Total} markers within the training population. Assuming the test population is genetically similar to the training population, the values of minimum M_{Total} and M_{Low} obtained from cross validation can then be used in imputation.

Table 3 Linkage disequilibrium (LD), accuracy of imputation (r_I), and accuracy of genomewide prediction (r_{MP}) with different numbers of single nucleotide polymorphism markers in a collection of 283 maize inbreds. Results are for a single partitioning of the maize inbreds into a training population and test population.

Type of population	M_{Low} ‡	M_{Total} §	Mean r^2 (LD) ¶	r_I #	r_{MP} † Ear height	Plant height	Anthesis date	Silking date	Oil	Protein	Starch
Non-imputed		96	0.03		0.26	0.34	0.50* ††	0.49*	0.26*	0.50*	0.44
		192	0.05		0.23*	0.36	0.55	0.53*	0.50	0.54	0.34*
		288	0.05		0.26	0.33	0.54*	0.52*	0.48	0.55	0.44
		384	0.06		0.35	0.41	0.64	0.61	0.49	0.58	0.44
		768	0.09		0.35	0.41	0.64	0.61	0.49	0.58	0.44
		1536	0.08		0.38	0.48	0.66	0.65	0.52	0.61	0.47
		3000	0.11		0.37	0.47	0.66	0.64	0.53	0.61	0.49
		6000	0.15		0.39	0.48	0.67	0.66	0.54	0.62	0.47
		9000	0.20		0.39	0.48	0.68	0.66	0.54	0.61	0.49
		12000	0.22		0.40	0.48	0.67	0.66	0.54	0.60	0.48
		14850	0.23		0.41	0.48	0.68	0.66	0.54	0.61	0.48
		30973	0.32		0.39	0.44	0.65	0.62	0.52	0.59	0.41
Imputed	96	1536		0.22	0.25*	0.40	0.31*	0.37*	0.34*	0.50*	0.41
	192	1536		0.41	0.35	0.47	0.52*	0.57	0.39*	0.52	0.44
	288	1536		0.51	0.36	0.47	0.62	0.62	0.41	0.59	0.47
	96	3000		0.14	0.15*	0.34	0.53*	0.52*	0.22*	0.45	0.33*
	192	3000		0.28	0.32	0.43	0.50*	0.48*	0.32*	0.54	0.47
	288	3000		0.38	0.33	0.41	0.57	0.58	0.36*	0.58	0.46
	96	6000		0.08	0.30	0.39	0.49*	0.50*	0.27*	0.37*	0.27*
	192	6000		0.19	0.32	0.40	0.54*	0.53*	0.29*	0.46*	0.37
	288	6000		0.27	0.26	0.34	0.50*	0.49*	0.26*	0.50*	0.44
	96	30973		0.04	-0.02*	0.16*	0.20*	0.26*	0.28*	0.37*	0.28*

	192	30973	0.07	0.26	0.36	0.36*	0.39*	0.43	0.47*	0.38
	288	30973	0.11	0.07*	0.18*	0.35*	0.37*	0.39*	0.51	0.42
	384	30973	0.14	0.28	0.34	0.44*	0.43*	0.43	0.58	0.44
	768	30973	0.26	0.38	0.44	0.54*	0.51*	0.42	0.58	0.47
	1536	30973	0.40	0.37	0.49	0.64	0.64	0.31*	0.58	0.37
	3000	30973	0.59	0.40	0.48	0.63	0.62	0.56	0.62	0.50
	6000	30973	0.75	0.41	0.49	0.68	0.66	0.56	0.62	0.48
g^2 ‡‡				0.74	0.34	0.86	0.84	0.64	0.79	0.58

† Correlation between genomewide predictions and observed performance

‡ For the imputed populations, one-third of inbreds had known genotypes at M_{Low} markers and the remaining $M_{Total} - M_{Low}$ markers were imputed.

§ Total number of markers

¶ Mean pairwise r^2 value between adjacent markers

Mean across inbreds in the test population and across 100 repeats with fastPHASE software

†† Significantly different (P = 0.05) from the highest observed r_{MP} for that trait based on a Fisher z transformation

‡‡ Ratio between the variance due to inbreds and the phenotypic variance on an entry-mean basis

Table 4 Linkage disequilibrium (LD), accuracy of imputation (r_I), and accuracy of genomewide prediction (r_{MP}) with different numbers of single nucleotide polymorphism markers in a collection of 860 barley inbreds. Results are for a single partitioning of the barley inbreds into a training population and test population.

Type of population	M_{Low} ‡	M_{Total} §	Mean r^2 (LD) ¶	r_I #	r_{MP} † Grain hardness	Grain width	Grain weight	Hull proportion	Polyphenol oxidase activity
Non-imputed		96	0.25		0.50* ††	0.82	0.85*	0.47*	0.63*
		192	0.28		0.55*	0.83	0.86	0.48	0.66*
		288	0.31		0.55*	0.83	0.86	0.48	0.66*
		384	0.30		0.60*	0.83	0.86	0.52	0.68*
		768	0.31		0.65	0.83	0.87	0.52	0.71*
		1536	0.28		0.68	0.83	0.87	0.54	0.72*
		2019	0.24		0.67	0.84	0.88	0.55	0.78
Imputed	96	768		0.73	0.50*	0.81	0.83*	0.41*	0.64*
	192	768		0.85	0.59*	0.83	0.86	0.47*	0.69*
	288	768		0.91	0.61*	0.83	0.87	0.49	0.70*
	96	1536		0.62	0.44*	0.79*	0.81*	0.41*	0.59*
	192	1536		0.78	0.61*	0.83	0.87	0.48	0.69*
	288	1536		0.84	0.63	0.83	0.87	0.51	0.71*
	384	1536		0.90	0.66	0.83	0.87	0.53	0.71*
	96	2019		0.52	0.42*	0.78*	0.78*	0.39*	0.55*
	192	2019		0.71	0.59*	0.83	0.86	0.48	0.70*
	288	2019		0.81	0.63	0.83	0.87	0.50	0.72*
	384	2019		0.86	0.66	0.84	0.88	0.53	0.73*
	768	2019		0.92	0.67	0.84	0.88	0.54	0.74
	1536	2019		0.96	0.67	0.84	0.88	0.54	0.75
g^2 ‡‡					0.50* ††	0.82	0.85*	0.47*	0.63*

† Correlation between genomewide predictions and observed performance

‡ For the imputed populations, one-third of inbreds had known genotypes at M_{Low} markers and the remaining $M_{Total} - M_{Low}$ markers were imputed.

§ Total number of markers

¶ Mean pairwise r^2 value between adjacent markers

Mean across inbreds in the test population and across 100 repeats with fastPHASE software

†† Significantly different ($P = 0.05$) from the highest observed r_{MP} for that trait based on a Fisher z transformation

‡‡ Ratio between the variance due to inbreds and the phenotypic variance on an entry-mean basis

Table 5 Mean accuracy of imputation (r_I) and genomewide prediction (r_{MP}) across five repeats for selected numbers of single nucleotide polymorphism markers in a collection of 283 maize inbreds.

Type of Population	M_{Low} ‡	M_{Total} §	r_I ¶	r_{MP} †						
				Ear height	Plant height	Anthesis date	Silking date	Oil	Protein	Starch
Non-imputed		192		0.33* #	0.44*	0.46*	0.47*	0.39*	0.52*	0.43*
		288		0.36*	0.43*	0.52*	0.51*	0.42	0.52*	0.46
		1536		0.45	0.52	0.64	0.63	0.46	0.57	0.47
		3000		0.44	0.53	0.65	0.64	0.47	0.58	0.48
		14850		0.47	0.54	0.68	0.66	0.46	0.58	0.48
Imputed	192	1536	0.52	0.42	0.49	0.54*	0.56*	0.36*	0.51*	0.46
	288	1536	0.44	0.40*	0.49	0.55*	0.56*	0.38*	0.56	0.48
	288	3000	0.39	0.40	0.49	0.58*	0.59*	0.42	0.56	0.48
LSD (P = 0.05)				0.07	0.06	0.06	0.05	0.06	0.05	0.03

† Mean correlation between genomewide predictions and observed performance

‡ For the imputed populations, one-third of inbreds had known genotypes at M_{Low} markers and the remaining $M_{Total} – M_{Low}$ markers were imputed.

§ Total number of markers

¶ Mean across inbreds in the test population and across 100 repeats with fastPHASE software

Significantly different (P = 0.05) from the highest observed r_{MP} for the trait based on the least significant difference (LSD)

Table 6 Mean accuracy of imputation (r_I) and genomewide prediction (r_{MP}) across five repeats for selected numbers of single nucleotide polymorphism markers in a collection of 860 barley inbreds.

				r_{MP} †				
Type of population	M_{Low} ‡	M_{Total} §	r_I ¶	Grain hardness	Grain width	Grain weight	Hull proportion	Polyphenol oxidase activity
Non-Imputed		288		0.57* #	0.82*	0.85*	0.49	0.69*
		384		0.60*	0.82*	0.86	0.51	0.70*
		768		0.65	0.83	0.86	0.51	0.73*
		1536		0.68	0.84	0.87	0.53	0.73*
		2019		0.68	0.84	0.87	0.53	0.79
Imputed	288	768	0.91	0.62*	0.82*	0.86	0.49	0.72*
	288	1536	0.86	0.63	0.83	0.86	0.49	0.71*
	384	1536	0.90	0.55*	0.83	0.86	0.46*	0.70*
LSD (P = 0.05)				0.05	0.01	0.01	0.04	0.02

† Mean correlation between genomewide predictions and observed performance

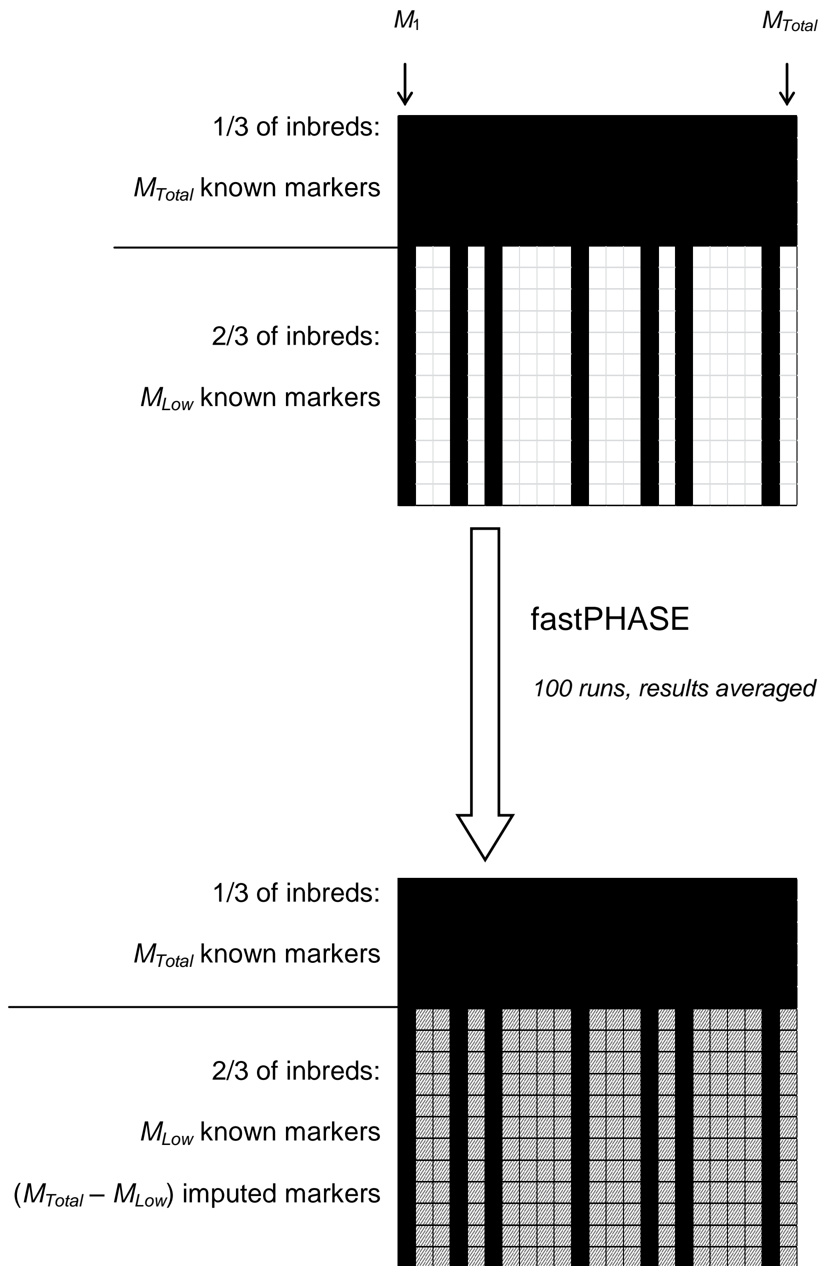
‡ For the imputed populations, one-third of inbreds had known genotypes at M_{Low} markers and the remaining $M_{Total} - M_{Low}$ markers were imputed.

§ Total number of markers

¶ Mean across inbreds in the test population and across 100 repeats with fastPHASE software

Significantly different (P = 0.05) from the highest observed r_{MP} for the trait based on the least significant difference (LSD)

Figure 2 Marker imputation scheme.



Chapter 3: Genomewide Selection to Introgress Semidwarf Corn Germplasm into U.S. Corn Belt Inbreds

Semidwarf corn (*Zea mays* L.) could potentially be grown in new areas of production or in alternative crop rotations. My objectives were to determine: (i) if genomewide selection is useful for the rapid improvement of an exotic × adapted cross; (ii) if genomewide selection is more effective than phenotypic backcrossing for a trait with major genes; and (iii) if the high grain yield of nondwarf corn can be combined with the reduced stature and adaptability to high plant population densities of semidwarf corn. I conducted four cycles of genomewide selection in two semidwarf × adapted crosses. Phenotypic backcrossing was also done until the BC_4 with selection for short plants. The accuracy of genomewide predictions in Cycle 0 was 0.67–0.70 for plant height and 0.57–0.70 for grain yield. Genomewide selection from Cycle 1 until Cycle 5 either maintained or improved upon the gains from phenotypic selection achieved in Cycle 1. Observed gains generally agreed with predicted gains. Cycle 5 did not always have the best mean performance. Antagonistic genetic correlations (0.74–0.75) made it difficult to select for short plants with high grain yield. Linkage disequilibrium between markers declined as selection progressed. Compared with phenotypic backcrossing, genomewide selection led to better mean performance and a higher proportion of exotic germplasm introgressed. To my knowledge, this is the first empirical study on genomewide selection to improve an exotic × adapted cross.

Introduction

Corn hybrids in the northern U.S. Corn Belt are typically >2m tall and have a relative maturity of 75–110d. COPOP1, an open-pollinated semidwarf corn population developed by researchers at Agriculture and Agri-Food Canada starting in the 1970s (L.M. Reid, pers. comm., 2004), grows less than 1m tall and has a relative maturity of about 62d (Schaefer et al., 2011). Whereas nondwarf corn hybrids are currently grown in rows spaced 0.76m apart and at plant population densities of 80,000–90,000 plants ha^{-1}, COPOP1 can be grown like small grains in narrow rows (e.g., 0.25m apart) and at plant population densities of up to 200,000 plants ha^{-1}.

Density-tolerant semidwarf corn adapted to the longer growing seasons of the upper Midwest has the potential to increase ecosystem services (Begna et al., 2001) and could be grown in new areas of production or in alternative crop rotations (Russnogle, 1998). As a form of exotic germplasm, COPOP1 can help introduce genetic diversity for sustained corn improvement (Hallauer et al., 1972; Goodman, 1999; Mikel and Dudley, 2006; Nelson et al., 2008). Initial studies at the University of Minnesota indicated COPOP1 is a potentially valuable source of dwarfing but, as an early maturing population, also has low grain yield (Schaefer et al., 2011).

Exotic germplasm has been introgressed into adapted germplasm by three methods: (i) prebreeding of the exotic population; (ii) selection in an exotic × adapted cross; and (iii) backcrossing of a few known genes or of a trait from the exotic donor into an adapted inbred. Theoretical and empirical studies have shown that 5–10 cycles of selection, which is equivalent to 10–20 years of breeding work, are generally required for prebreeding (Hallauer et al., 1972; Crossa and Gardner, 1987; Goodman, 1999). Several

cycles of selection are likewise needed for prebreeding in an exotic × adapted cross (Albrecht and Dudley, 1987; Goodman, 1999; Gousenard et al., 1996; Jumbo et al., 2011). Backcrossing a few known genes or a trait to the BC_5 or later generation introgresses only a limited proportion of exotic germplasm and still often results in yield drag (Hoffbeck et al., 1995; Šimić et al., 2003).

Simulation studies have shown that genomewide selection (or genomic selection; Meuwissen et al., 2001) may be an efficient means to introgress exotic germplasm into an adapted inbred (Bernardo, 2009). In particular, genomewide selection could shorten the time required for prebreeding in an exotic × adapted cross from 10–20 years to less than four (Bernardo, 2009). This reduction in time is achieved by the use of molecular markers to select for complex traits in a year-round nursery or greenhouse, where marker data are meaningful but where phenotypic data are not. Furthermore, because genomewide selection acts on the entire genome, it can be used to select for favorable alleles already present in the adapted germplasm (Heffner et al., 2009).

No empirical studies have been reported on the use of genomewide selection to introgress exotic germplasm into adapted inbreds. My objectives were to determine: (i) if, as indicated by theoretical studies (Bernardo, 2009), genomewide selection is useful for the rapid improvement of an exotic × adapted cross for multiple traits; (ii) if genomewide selection is more effective than phenotypic backcrossing for a trait with major genes (dwarf stature); and (iii) if the high grain yield of nondwarf corn can be combined with the reduced stature and adaptability to high plant population densities of semidwarf corn.

Materials and methods

Overview

In genomewide selection, F_3 families derived from each of two semidwarf × adapted corn crosses served as Cycle 0. Testcrosses of the Cycle 0 F_3 families were evaluated in yield trials, and phenotypic selection based on a Yield Index and plant height was done to create Cycle 1. Four additional cycles of selection (until Cycle 5) were conducted based on genomewide markers alone. As an alternative to genomewide selection, phenotypic backcrossing was done until the BC_4S_1 or BC_4S_2 stage with selection for short plants. At the end of genomewide selection and phenotypic backcrossing, field trials of bulks from each cycle were conducted to measure responses to selection.

Germplasm

I studied two semidwarf S_1 lines (DC1 and DC2) and two nondwarf inbreds (LH74 and PHG50) whose Plant Variety Protection certificates have expired. DC1 and DC2 were developed by one generation of selfing from COPOP1 and were selected based on their performance in yield trials at the University of Minnesota (Schaefer et al., 2011). LH74 was developed by Monsanto and belongs to the Iowa Stiff Stalk Synthetic (BSSS) heterotic group (MBS, 2003) whereas PHG50 is an Iodent inbred developed by Pioneer Hi-Bred (Mikel and Dudley, 2006). I considered two exotic × adapted crosses: DC1 × LH74 and DC1 × PHG50. DC2 was used as the tester for both crosses. As described later, I attempted to breed DC1 × LH74 for later maturity (i.e., maturity of LH74) and DC1 × PHG50 for earlier maturity (i.e., maturity of DC1).

Cycle 0 and Parental Screens

A single DC1 plant was selfed to form an S_2 line and was also crossed to both LH74 and PHG50 in January–April 2009 at Molokai, HI. The resulting F_1 crosses were

selfed to form the F_2 generation in June–September 2009 at the University of Minnesota Saint Paul campus. Random F_2 plants were selfed in September–December 2009 at Molokai to form 152 (DC1 × LH74) F_3 families and 163 (DC1 × PHG50) F_3 families.

Parental screening with 2000 single nucleotide polymorphism (SNP) markers was conducted as follows. The S_2 progeny of the DC1 plant used for crossing at Molokai were grown, along with LH74 and PHG50, at the University of Minnesota Plant Growth Facilities Greenhouses in October–November 2009. Leaf punches were taken from 2 week old leaves of 10 S_2 plants of DC1 and two plants of the adapted parents, and the leaf tissue was desiccated and shipped to DNA LandMarks (Saint-Jean-sur-Richelieu, Quebec) where DNA was extracted and genotyping was done. In addition, leaf punches were obtained from the individual F_2 plants that were being selfed at Molokai. A bulk of the DNA from the (DC1 × LH74) F_2 plants and a bulk of DNA from the (DC1 × PHG50) F_2 plants were included in parental screening to ensure that SNPs polymorphic between DC1 and the adapted parents were also polymorphic in the F_2 populations. A set of 288 polymorphic markers were chosen for DC1 × LH74 and a separate set of 288 polymorphic markers were chosen for DC1 × PHG50. Priority was given to markers that were fixed in DC1 and to markers that were evenly distributed across the genome. The two sets of SNP markers were then used in subsequent genomewide selection in the respective crosses. Because DC1 was still segregating, 151 markers used for the DC1 × LH74 cross were heterozygous in DC1, and 181 markers used for the DC1 × PHG50 cross were heterozygous in DC1. The DNA from the individual F_2 plants that were selfed to form each of the 152 (DC1 × LH74) F_3 families and 163 (DC1 × PHG50) F_3 families was then analyzed for the 288 SNPs within each cross.

Cycle 0 testcrosses were made at Molokai in January–April 2010 by crossing a bulk of pollen from 10–12 plants of each F_3 family to the DC2 tester. The Cycle 0 testcrosses were evaluated in 2010 for grain yield (Mg ha^{-1}, adjusted to 155 g H_2O kg^{-1}), grain moisture (g kg^{-1}), root lodging (percentage of plants leaning from the vertical at an angle of 45° or more), stalk lodging (percentage of plants with stalks broken below the ear), and plant height (distance, in cm, from the soil surface to the tip of the corn tassel). The Cycle 0 yield trials were conducted at two locations at the University of Minnesota Southwestern Research and Outreach Center at Lamberton and at two locations at the University of Minnesota Southern Research and Outreach Center at Waseca. The two locations within each Research and Outreach Center had different soil types (Normania loam and Ves loam at Lamberton and Webster clay loam and Nicollet clay loam at Waseca) and were planted approximately 10d apart. An augmented randomized complete block design was used with the two crosses evaluated in separate but adjacent blocks in each location. Plots were 6.1m long with five rows per plot and 0.51m between rows. Data were collected only from the three center rows. Plant population density at the different locations ranged from 105,000–129,000 plants ha^{-1}.

Cycle 0 data were analyzed using PROC UNIVARIATE of the SAS software version 9.2 for Windows 7 (Cary, NC) to obtain means and standard errors of the mean for each trait. Testcross genetic variance (V_G) and heritability (h^2) of each trait were estimated on an entry-mean basis from ANOVA. The significance of V_G was tested using an F-test and the significance of h^2 was tested by constructing a 90% confidence interval, with h^2 considered significant if the interval did not contain zero (Knapp, 1985).

Genomewide marker effects for each cross were estimated based on the performance of the Cycle 0 testcrosses for each trait and SNP data on the progenitor F_2 plants. Marker effects were estimated by ridge regression-best linear unbiased prediction (RR-BLUP) in SAS PROC IML (Meuwissen et al., 2001). The accuracy of genomewide predictions was estimated as the correlation (r_{MP}) between testcross genotypic values predicted from genomewide marker effects and phenotypic values of the testcrosses for each trait. Specifically, r_{MP} was calculated from five-fold cross validation (Utz et al., 2000). The r_{MP} was also calculated for the Yield Index (described below), for which the approximate V_G and V_P were obtained as the variance of a linear function while ignoring any covariance between traits in the Yield Index.

For plant height, QTL mapping was done by Haley-Knott regression (Knott and Haley, 1992) using the package R/QTL version 1.22-21 for R version 2.12.2 for Windows 7 (Broman et al., 2003). The per-marker significance threshold that corresponded to an experiment-wise Type I error rate of $P = 0.05$ was determined from 300 Haley-Knott permutations. Best linear unbiased estimates of the testcross mean of each F_3 family for grain yield and plant height were calculated using SAS PROC GLM, and phenotypic correlations between plant height and grain yield were calculated as the Pearson correlation coefficient between the entry means for both traits. Genotypic correlations between plant height and grain yield were calculated as the Pearson correlation coefficient between the RR-BLUP marker effects for the two traits (Ziyomo, 2012).

Multi-trait Selection Indices

A selection index was used to combine information from multiple traits into a single selectable criterion. First, a Yield Index was constructed for each cross as *Y* = (grain yield in Mg ha^{-1}) – 0.028 (grain moisture in g H_2O kg^{-1}) – 0.059 (stalk lodging percentage) – 0.036 (root lodging percentage), where the weights were the retrospective-index weights (Dickerson et al., 1954) that have been used by a team of experienced commercial corn breeders (Bernardo, 1991). Given that selection for moisture, as an indicator of maturity, was not done in the DC1 × LH74 cross, the Yield Index in this cross did not include grain moisture. The testcrosses were then ranked according to the Yield Index values with a rank of 1 for the highest.

Second, the testcrosses were ranked according to plant height with a rank of 1 for the shortest. The final selection index, which combined the Yield Index and plant height, was calculated as the sum of ranks for the Yield Index and for plant height. The selection index therefore gave equal weights to grain yield and agronomic traits as a whole and to plant height.

Phenotypic Selection of Cycle 0 Individuals to Form Cycle 1

Previous research has shown that the response at the end of the entire selection process is greater when selection in Cycle 0 is based solely on phenotypic data, than if selection in Cycle 0 is based on both phenotypic data and marker data (Bernardo and Yu, 2007). Selection in Cycle 0 was therefore based on phenotypic data only. Testcrosses of the F_3 families in each cross were ranked according to their selection index scores. In each cross, the 10 F_3 families with the best (lowest) selection index scores were intermated to form Cycle 1. Each F_3 family was crossed with a balanced bulk from the nine other families. Equal numbers of the resulting seeds were bulked from each F_3

family. These crosses were made at the University of Minnesota Plant Growth Facilities Greenhouses in October 2010–January 2011.

Genomewide Selection in Cycles 1 to 5

Genomewide selection in Cycle 1 to 5 was performed based on the RR-BLUP prediction equations developed in Cycle 0 for the DC1 × LH74 and DC1 × PHG50 crosses. For each of the two crosses, 192 individual plants in each cycle of selection were genotyped prior to flowering by collecting leaf samples at the seedling stage. In each cross, the set of 288 SNP markers used in Cycle 0 were used in subsequent cycles. Genotypic values were predicted for each cycle of selection as $\hat{\mathbf{y}}_\mathbf{p} = \mu + \mathbf{M}\hat{\mathbf{g}}$ where $\hat{\mathbf{y}}_\mathbf{p}$ was a 192 × 1 vector of predicted genotypic values of the plants in a given cycle of genomewide selection; μ was the grand mean; **M** was an 192 × 288 incidence matrix of coded genotypes (1 and –1 for the contrasting homozygotes and 0 for the heterozygote); and $\hat{\mathbf{g}}$ was a 288 × 1 vector of RR-BLUP marker effects estimated from Cycle 0 for each trait in each cross (Meuwissen et al., 2001).

In each cross and in each cycle, grain yield, moisture, stalk lodging, root lodging, and plant height of each of the 192 plants was predicted. These predicted values were then used to construct the same selection index used in Cycle 0. In each cross, the 10 plants with the best (lowest) selection index were intermated by a standard chain-crossing method in which the earliest selected plant was crossed with the second-earliest selected plant; the second earliest selected plant was crossed to the third-earliest selected plant; and so on, until the latest flowering selected plant was crossed to the earliest flowering selected plant. The next cycle was formed by bulking equal numbers of seeds from the resulting ears. The numbers of plants evaluated and selected were constant across cycles.

Phenotypic Backcrossing

To compare genomewide selection with phenotypic backcrossing for a trait with major genes (dwarf stature), F_1 plants from each cross were backcrossed to LH74 or PHG50 in summer 2009 at Saint Paul. The resulting BC_1 plants were selfed at Molokai in September–December 2009. In January–April 2010, 75 BC_1S_1 plants from each cross were grown at Molokai and the shortest plant was backcrossed to the recurrent parent to form the BC_2 generation. Starting at the BC_2 stage, all crosses were made at greenhouses at the University of Minnesota Plant Growth Facilities. The BC_2 plants were selfed to form the BC_2S_1 generation, and the procedures used in the first backcross generation were repeated until BC_4S_1 plants were generated. In the LH74 cross, BC_4S_1 plants were selfed to increase seed. The DC1 × PHG50 cross had one less stage of selfing due to a pollination failure that required one generation to be repeated.

Progress from Selection

In each cross, about 80 plants from each cycle of genomewide selection and from the last generation of phenotypic backcrossing were testcrossed to DC2 and to an adapted tester inbred from Monsanto (LH227). Testcrosses were made at Molokai and at the greenhouses at the University of Minnesota Plant Growth Facilities. Seeds were bulked to form a total of 27 entries: (i) Cycles 0–5 of DC1 × LH74, testcrossed to DC2; (ii) Cycles 0–5 of DC1 × LH74, testcrossed to LH227; (iii) Cycles 0–5 of DC1 × PHG50, testcrossed to DC2; (iv) Cycles 0–5 of DC1 × PHG50, testcrossed to LH227; (vi) last generation of phenotypic backcrossing in DC1 × LH74 testcrossed to DC2; and (vii) last generation of phenotypic backcrossing in DC1 × PHG50 testcrossed to DC2 and LH227.

The last generation of phenotypic backcrossing in DC1 × LH74 was also testcrossed to LH227 but not enough seeds were produced in time for yield trials.

The 27 testcrosses were evaluated, along with three commercial check hybrids, at six locations in 2012. Four of the locations were the same at in 2010: two locations at University of Minnesota Southwestern Research and Outreach Center at Lamberton and two locations the University of Minnesota Southern Research and Outreach Center at Waseca. Testcrosses were also evaluated at two locations at the University of Minnesota Saint Paul campus. The two locations at each outreach center and on campus had different soil types when possible (Normania loam and Ves loam at Lamberton, Nicollet clay loam and Webster clay loam at Waseca, both Waukegan silk loam at Saint Paul) and were planted approximately 14d apart. A randomized complete block design with three replications was used. Plots were 4.58m long at Lamberton and Waseca and 3.66m long at Saint Paul. All locations had five rows per plot and 0.51m row spacing. Data were collected only from the three center rows. Plant population density across locations was 115,000–152,000 plants ha^{-1}. Testcrosses were evaluated for the same traits as the Cycle 0 testcrosses in 2010. Due to difficulties with the combine, grain yield and moisture data from the later planting date at Lamberton were excluded from analysis.

To account for some missing plots, best linear unbiased estimates of the performance of each entry were calculated using SAS PROC MIXED. Least significant differences ($P = 0.05$) were calculated based on the mean standard error of the differences among all pairwise comparisons between entries (Vargas et al. 2012). Each cross × tester combination was analyzed separately.

Given that genomewide selection started in Cycle 1, the observed response to genomewide selection was calculated as the difference between the mean at each later cycle and the mean at Cycle 1. Confidence intervals (P = 0.95) on the observed responses to genomewide selection were obtained. Linear contrasts were tested for significance in PROC MIXED. For each cross × tester combination, I tested the following contrasts: (i) linear effect from Cycle 1 to Cycle 5; (ii) Cycle 5 vs. Cycle 1; and (iii) Cycle 1 vs. Cycle 0. In addition, I tested several contrasts of phenotypic backcrossing versus different cycles or combinations of cycles of genomewide selection.

Predicted Response and Population Genetics of Cycle 0 to Cycle 4 Entries

For each trait, the predicted mean performance in each cycle (with DC2 as the tester) was obtained by multiplying the mean allele frequency (across the 192 individuals in a given cycle) at a SNP marker by the RR-BLUP effect for that marker, summing the products across all 288 SNP markers, and adding the Cycle 0 grand mean. Mean allele frequencies in Cycle 5 were calculated from the allele frequencies in the 10 selected plants from Cycle 4. Predicted responses to genomewide selection were calculated as the difference between the predicted mean at each later cycle and the predicted mean at Cycle 1.

For the different cycles of genomewide selection, I examined the mean change in SNP allele frequency from the previous cycle, the inbreeding coefficient, and linkage disequilibrium (LD). The mean change in SNP allele frequency was calculated as the mean absolute value of the change in allele frequency across all markers. The inbreeding coefficients were calculated as $1 - H_{Cn}/H_{C0}$ where H_{Cn} was the mean marker heterozygosity in Cycle n and H_{C0} was the mean marker heterozygosity in Cycle 0. The

LD was calculated using Haploview Version 4.2 (Barrett et al., 2005) as the mean r^2 value between adjacent SNP markers. Only SNP markers with known genetic positions (157 markers for the DC1 × LH74 cross and 121 markers for the DC1 × PHG50 cross) were used in calculating LD.

Results and Discussion

Performance of Cycle 0

The Cycle 0 testcrosses had moderate plant height, moderate grain yield, and low root and stalk lodging (Table 7). Plant height and grain moisture in Cycle 0 were lower in the DC1 × PHG50 cross than in the DC1 × LH74 cross. All traits had significant V_G and h^2 in each cross, with the h^2 for grain yield being 0.92 in the DC1 × PHG50 cross and 0.91 in the DC1 × LH74 cross). The high h^2 for grain yield may have reflected the strong correlation between plant height, which differed widely among the testcrosses, and grain yield. The phenotypic correlations between plant height and grain yield were 0.80 for DC1 × LH74 and 0.82 for DC1 × PHG50, whereas the genotypic correlations were 0.74 for DC1 × LH74 and 0.75 for DC1 × PHG50. Testcross means for plant height ranged from 113 to 221cm in the DC1 × PHG50 cross and from 114 to 233cm in the DC1 × LH74 cross.

For the Yield Index, the accuracy of genomewide prediction (r_{MP}) was 0.41 in the DC1 × LH74 cross and 0.46 in the DC1 × PHG50 cross (Table 7). The r_{MP} was >0.50 for the two main traits, plant height and grain yield, in the selection index. The r_{MP} was lower for root lodging, stalk lodging, and moisture. The higher r_{MP} values for grain yield and plant height and lower r_{MP} values for the three other traits were consistent with their h^2. In general, r_{MP} was higher in the DC1 × PHG50 cross than the DC1 × LH74 cross. For

grain yield, r_{MP} was 0.70 for the DC1 × PHG50 cross compared to 0.57 for the DC1 × LH74 cross. Overall, the significant V_G and the moderate to high r_{MP} for the selection index and for the component traits suggested that genomewide selection would be effective in each cross.

One SNP interval (pair of adjacent markers) was found significant for plant height in each cross (Fig. 3). In the DC1 × LH74 cross, the significant interval had a LOD score of 33.0 and was at the 178 cM position on Chromosome 1. The QTL allele from DC1 had a per-copy effect of –15.8cm and the QTL explained 63.5% of the variation for plant height in the DC1 × LH74 cross. In the DC1 × PHG50 cross, the significant marker had a LOD score of 36.4 and was at the 179 cM position on Chromosome 1. The QTL allele from DC1 had a per-copy effect of –15.1 cm and the QTL explained 63.4% of the variation for plant height in the DC1 × PHG50 cross. The significant markers were those closest to the known position of a major dwarfing gene, reduced stature 1 (*rd*1; Beavis et al., 1991), confirming that *rd*1 is present in DC1.

Response to Genomewide Selection

In a previous study, one cycle of testcross phenotypic selection in COPOP1 did not lead to a significant improvement in grain yield, plant height, and other agronomic traits (Schaefer et al., 2011). In the DC1 × LH74 cross in this study, plant height increased significantly ($P = 0.05$) due to phenotypic selection from Cycle 0 to Cycle 1 but the Yield Index did not change. In the DC1 × PHG50 cross, both plant height and the Yield Index increased significantly from Cycle 0 to Cycle 1. Among the component traits of the Yield Index, only grain yield in the DC1 × PHG50 cross changed significantly from Cycle 0 to Cycle 1.

Genomewide selection, conducted from Cycle 1 until Cycle 5, either maintained or improved upon the gains from phenotypic selection achieved in Cycle 1. In the DC1 × LH74 cross, a significant linear contrast (from Cycle 1 to 5) indicated a per-cycle reduction in plant height due to genomewide selection. The corresponding linear contrast for the Yield Index was not significant, although the per-cycle reduction in grain yield was significant. Moisture was not included in the Yield Index in the DC1 × LH74 cross, and the linear contrasts for root and stalk lodging from Cycle 1 to 5 were nonsignificant. In the DC1 × PHG50 cross, genomewide selection did not result in significant linear changes from Cycle 1 to 5 in plant height, the Yield Index, or the component traits.

The predicted and observed gains from genomewide selection were similar for most traits across all cycles of selection (Fig. 4). The largest differences, for which the 95% confidence interval on the observed response did not overlap with the predicted response, were for plant height in Cycle 5 in the DC1 × PHG50 cross, grain yield in Cycle 5 of the DC1 × LH74 cross, and grain yield in Cycles 4 and 5 of the DC1 × PHG50 cross. Overall, predicted and observed gains were most similar for plant height, especially in the DC1 × LH74 cross. The markers for *rd*1 were fixed by Cycle 1 of the DC1 × PHG50 cross but were not fixed until Cycle 5 of the DC1 × LH74. The continued predicted response after Cycle 1 in the DC1 × PHG50 cross therefore reflected predicted gains for selection at plant height QTL other than *rd*1.

For the Yield Index, the predicted and observed gains were not significantly different in any cycle. The small predicted gains helped explain why I saw little significant progress for the Yield Index. Predicted gains for root lodging and stalk

lodging were small, ranging from zero to –3% (results not shown) and were not significantly different from the observed gains in either cross.

Neither cross achieved maximum gain for all traits in Cycle 5. In the DC1 × LH74 cross, plant height was lowest in Cycle 5 and significantly different from Cycle 1 (P = 0.05); the Yield Index was the highest in Cycle 2 but was not significantly different from Cycle 1. The selection index (i.e., combination of plant height and Yield Index) was best in Cycle 2. In the DC1 × PHG50 cross, plant height was lowest in Cycle 4 but was not significantly different from Cycle 1. The Yield Index was highest in Cycle 3 but not significantly different from Cycle 1; the Yield Index was highest in Cycle 3 but was not significantly different from Cycle 1. The selection index was the equally high in Cycles 3, 4 and 5.

I believe that the lack of sustained response to genomewide selection was due to three factors: (i) counteracting effects of selection for plant height versus grain yield; (ii) inherent variability in responses to marker-based selection; and (iii) decline in LD during selection. The strong association between tall plants and high grain yield (genetic correlations of 0.74–0.75) made it very challenging to select for a combination of low plant height and high grain yield. The antagonistic correlation between plant height and grain yield was also reflected in the predicted responses to genomewide selection (Fig. 4). Poor grain yields are a known drawback of semidwarf corn, with hybrids containing the dwarfing gene *rd*1 yielding about half of hybrids without the dwarfing gene (Dijak et al., 1999; Schaefer et al., 2011). It was therefore notable that genomewide selection was able to decrease plant height significantly while maintaining the Yield Index in the DC1 × LH74 cross.

The responses to genomewide selection, marker-assisted recurrent selection (based on significant markers only; Bernardo and Yu, 2007), and recurrent phenotypic selection have been known to vary over several cycles. Although simulation experiments for both marker-assisted recurrent selection and genomewide selection suggested that gain could be achieved over several cycles (Bernardo and Yu, 2007; Bernardo, 2009; Mayor and Bernardo, 2009), previous empirical experiments have shown erratic gains from multiple cycles of selection based on markers (Johnson, 2004; Massman et al., 2012). Studies of recurrent phenotypic selection where remnant seed from each cycle is tested also sometimes show erratic gains per cycle (Leng, 1961; Dudley, 2007). Given these results, the breeder should evaluate seeds from all cycles of genomewide selection (just as they would for phenotypic selection) by field testing.

The LD, measured as the mean r^2 between adjacent markers, declined rapidly (Table 9). The DC1 × PHG50 cross in particular showed a steep decrease in LD between Cycle 0 ($r^2 = 0.67$) and Cycle 1 ($r^2 = 0.38$); in contrast, the DC1 × LH74 cross showed a smaller decline between Cycle 0 ($r^2 = 0.65$) and Cycle 1 ($r^2 = 0.50$). This result could help explain why linear decreases in plant height were achieved until Cycle 5 in DC1 × LH74. By Cycle 4, the mean r^2 had declined to 0.26 in the DC1 × LH74 cross and 0.31 in the DC1 × PHG50 cross. While previous research has suggested that r^2 as low as 0.10–0.20 is sufficient for accurate genomewide predictions (Calus et al. 2008; Hayes et al. 2009), this previous research involved updating the genomewide prediction model yearly. A previous study in corn did show that genomewide selection can be effective until Cycle 3, but gains for some traits were not significant and the mean r^2 declined only to 0.40 (Massman et al., 2012).

Correlated Responses with a Different Tester

The use of a nondwarf tester (LH227) instead of a semidwarf tester (DC2) led to taller plants, higher grain yields, and more root lodging in both the DC1 × LH74 and DC1 × PHG50 crosses (Table 10). The gains from Cycles 1 to 5 with LH227 as the tester were different from the gains with DC2 as the tester with most of the differences being in grain yield. As previously mentioned, plant height decreased significantly and the Yield Index was constant from Cycles 1 to 5 in the DC1 × LH74 cross with DC2 as the tester. With LH227 as the tester, plant height was constant, the Yield Index and grain yield decreased linearly, and moisture increased linearly. Pairwise comparison of Cycle 1 versus Cycle 5 with LH227 as the tester showed a significant decrease in the Yield Index, decrease in grain yield, and increase in moisture. In the DC1 × PHG50 cross, whereas plant height and the Yield Index did not change significantly from Cycle 1 to 5 with DC2 as the tester, plant height still did not change but the Yield Index decreased linearly from Cycles 1 to 5 with LH227 as the tester. Pairwise comparison of Cycle 1 versus Cycle 5 with LH227 as the tester showed a significant decrease in the Yield Index with no changes in other traits.

While the responses to selection with LH227 were different than with DC2, most of the differences were for grain yield. Response for root and stalk lodging was the same with either tester. The different responses with the two testers for the Yield Index and grain yield were most likely due to (i) strong line × tester interaction for grain yield and (ii) a short growing season that put longer-season hybrids at a disadvantage. The influence of the tester is known to increase as the importance of non-additive genetic effects (e.g., dominance) increases (Rawlings and Thompson, 1962), and dominance

effects are more important for grain yield than for other agronomic traits in corn (Hallauer and Miranda, 1981, p. 116, 122). The 2012 growing season was short due to late planting and early harvest. This short season could have caused the grain yield of later cycles of selection testcrossed to LH227, which tended to have higher moisture and therefore later maturity, to be lower than they would otherwise have been.

Genomewide Selection vs. Phenotypic Backcrossing

An alternative to genomewide selection for introgressing semidwarf germplasm is to simply backcross, with phenotypic selection, the dwarfing trait from DC1 into LH74 and PHG50. In the DC1 × LH74 cross, the gain from phenotypic backcrossing (Cycle 0 to the BC_4 generation) was not significantly different from the gain from Cycle 0 to 5 for any trait. However, Cycle 2 of genomewide selection (for which the selection index was highest) had a significantly higher Yield Index and grain yield compared to phenotypic backcrossing, with no differences in plant height or the other traits. In the DC1 × PHG50 cross, the Yield Index was significantly higher and stalk lodging was significantly lower in Cycle 5 of genomewide selection than with phenotypic backcrossing, with no significant differences in the other traits. Because no linear contrasts were significant from Cycles 1 and 5 of genomewide selection, I also compared the gain from phenotypic backcrossing with the mean gains from Cycles 2 to 5. This contrast indicated that genomewide selection resulted in a significantly higher Yield Index, higher grain yield, lower moisture, and less stalk lodging compared to phenotypic backcrossing.

Overall, these results indicated that the mean performance across different traits was better with genomewide selection than with phenotypic backcrossing (Table 8). Genomewide selection was also more effective at incorporating exotic germplasm,

resulted in less inbreeding, and was faster than phenotypic backcrossing. At Cycle 5, the mean expected frequency (based on the genotypes of the selected Cycle 4 plants) of SNP marker alleles unique to DC1 was 0.41 in the DC1 × LH74 cross and 0.39 in the DC1 × PHG50 cross (Table 9). This proportion of exotic germplasm is similar to that retained in long term phenotypic selection experiments to introgress tropical material into temperate lines (Lewis et al., 2003). In contrast, the expected frequency of the exotic allele at the BC_4S_1 stage of phenotypic backcrossing is 0.03. General broadening of the genetic base, not just introducing new traits of interest, is an important goal of exotic × adapted crosses to support long term genetic gain (Goodman, 1999). Furthermore, an analysis of the marker effects for each trait revealed that beneficial marker alleles for each trait came approximately equally from the adapted (LH74 or PHG50) and exotic (DC1) parents. Therefore, beyond simply broadening the genetic base, incorporation of DC1 alleles should improve the traits of interest. The higher proportion of exotic germplasm with genomewide selection was consistent with less inbreeding with genomewide selection than phenotypic backcrossing (Table 9). The inbreeding coefficient for Cycle 4 was 0.30 in the DC1 × LH74 cross and 0.36 in the DC1 × PHG50 cross. Although inbreeding increased with each generation of selection, inbreeding at Cycle 5 was undoubtedly much less than the expected inbreeding coefficient of 0.94 at the BC_4 generation. The selection intensity was lower for genomewide selection, in which 10 plants out of 192 were selected in each cycle, than in phenotypic backcrossing, in which a single plant out of 96 was selected in each generation. The higher selection intensity might suggest that we would see greater gain the in phenotypic backcrossing population than in the genomewide selection population; however, we saw equivalent gain for plant height (the

only trait selected for in both populations) in spite of the stronger selection pressure in the phenotypic backcrossing populations.

Each cycle of genomewide selection was also faster than phenotypic backcrossing in this study: with genomewide selection all of the material could be planted simultaneously and all lines flowered at approximately the same time. With phenotypic backcrossing, multiple delay plantings of LH74 and PHG50 were needed for flowering to occur simultaneously between the backcross plants and recurrent parent. Furthermore, due to a higher percentage the genome from the adapted parent, the plants in phenotypic backcrossing tended to flower later than plants in genomewide selection. Phenotypic backcrossing took approximately two weeks longer per cycle than genomewide selection.

Prospects for Developing Semidwarf Hybrids Adapted to the U.S. Corn Belt

Nondwarf hybrids often have problems with stalk and root lodging at very high plant population densities (Sangoi, 2001). For example, one of the three check hybrids in this study [Pioneer 39F59, 77d to relative maturity (DRM)] had 35% root lodging and 10% stalk lodging (Table 8). Dropped ears and barrenness are also known problems at high plant population densities (Cox, 1996), but data were not collected for these traits because they were not an issue in this study. While many of the semidwarf testcrosses had low root and stalk lodging, the major challenge of the semidwarf corn was low grain yield, with the COPOP1 donor yielding less than half as much as nondwarf check hybrids in previous studies (Schaefer et al., 2011). As previously mentioned, other studies of corn hybrids with the dwarfing gene *rd*1 have also shown that semidwarf corn yields about half as much as nondwarf hybrids (Dijak et al., 1999).

The antagonistic genetic correlation between plant height and grain yield would impede progress in breeding for both traits regardless of whether breeding is by phenotypic selection or genomewide selection. Nevertheless, my results in the DC1 × LH74 cross suggest that genomewide selection may enable gains for plant height while holding grain yield and agronomic traits constant. Furthermore, the best cycles of genomewide selection (Cycle 2 of the DC1 × LH74 cross and Cycle 5 of the DC1 × PHG50 cross, with DC2 as the semidwarf tester) seemed to begin to combine the lodging tolerance of DC1 and DC2 (Schaefer et al., 2011) with the grain yield of nondwarf North American germplasm (Table 8). Cycle 2 of the DC1 × LH74 cross yielded 6.5 Mg ha^{-1} and had relatively little root lodging (7%) and stalk lodging (5%). Cycle 5 of the PHG50 cross yielded 5.7 Mg ha^{-1} and had relatively little root lodging (4%) and stalk lodging (0%). The check hybrid with the earliest relative maturity (Syngenta NO3-D8, 68 DRM) yielded 6.1 Mg ha^{-1} and had relatively high root lodging (17%) and stalk lodging (6%). None of the genomewide selection cycles were competitive in grain yield with the later DRM checks, BIXXIO (74 DRM) and 39F59 (77 DRM), which yielded 7.5 and 7.6 Mg ha^{-1}. Grain moisture was higher in both Cycle 2 of DC1 × LH74 (188 g kg^{-1}) and Cycle 5 of DC1 × PHG50 (168 g kg^{-1}) than in the commercial check hybrids (122–143 g kg^{-1}) and would need to be substantially improved. The commercial hybrids were the result of many years and environments of selection, and further cycles of selection in semidwarf × adapted crosses such as DC1 × LH74 or DC1 × PHG50 as well as improvement of the semidwarf tester, DC2, are needed to more fully assess the prospects of developing semidwarf hybrids for the U.S. Corn Belt.

Overall, I find these results promising for the future use of genomewide selection. Although I did not achieve consistent gain for all traits over all cycles of selection, I did produce lines that were superior to phenotypic backcrossing and was able to decrease plant height while holding the antagonistically correlated trait grain yield constant. I would anticipate that results would be even better for more common breeding situations where the traits under selection are not antagonistically correlated.

Table 7 Trait means, testcross genetic variance (V_G), entry-mean heritability (h^2) and correlation between predicted and observed performance (r_{MP}) in Cycle 0 testcrosses (to DC2) of 152 (DC1 × LH74) F_3 families and 163 (DC1 × PHG50) F_3 families. Tests were conducted in 2009 at four locations in Minnesota.

	DC1 × LH74					DC1 × PHG50				
Trait	Mean	SE†	V_G	h^2‡	r_{MP}§	Mean	SE	V_G	h^2	r_{MP}
Plant height (cm)	183	1	518*	0.95*	0.67	169	1	565*	0.92*	0.73
Stalk lodging (%)	3.1	0.2	7.4*	0.61*	0.40	1.9	0.2	2.9*	0.40*	0.31
Root lodging (%)	5.1	0.3	11.5*	0.63*	0.23	4.7	0.3	12.0*	0.63*	0.50
Moisture (g kg^{-1})	226	1	185*	0.61*	0.43	209	2	416*	0.73*	0.62
Grain yield (Mg ha^{-1})	6.32	0.06	1.25*	0.91*	0.57	5.30	0.06	1.30*	0.92*	0.70
Yield Index	5.47	0.11	1.41*	0.90*	0.46	–0.96	0.08	1.29*	0.89*	0.41

† Standard error of the mean

‡ Significance of h^2 was based on a 90% confidence interval

§ Mean correlation between predicted genotypic value and observed phenotypic value in five-fold cross-validation

Table 8 Testcross performance (with DC2 as the tester) of different cycles of genomewide selection and of progeny from phenotypic backcrossing in the DC1 × LH74 and DC1 × PHG50 corn crosses. Tests were conducted in 2012 at six locations in Minnesota.

Cross and cycle	Plant height (cm)	Plant height rank†	Yield Index‡	Yield Index rank	Selection Index score	Grain yield (Mg ha^{-1})	Moisture (g kg^{-1})	Root lodging (%)	Stalk lodging (%)
DC1 × LH74									
Cycle 0	191	2	4.5	6	8	5.5	171	10	5
Cycle 1	215	6	5.2	4	10	6.0	178	12	5
Cycle 2	203	3	6.0	1	4	6.5	188	7	3
Cycle 3	208	5	5.2	3	8	5.9	174	10	4
Cycle 4	206	4	5.9	2	6	6.4	181	11	3
Cycle 5	177	1	4.8	5	6	5.2	184	7	4
Phenotypic backcrossing	197		4.2			4.8	195	8	4
LSD§	20		1.0			0.8	33	8	5
DC1 × PHG50									
Cycle 0	152	1	–1.8	6	7	4.2	181	11	5
Cycle 1	187	5	0.6	5	10	5.9	172	9	2
Cycle 2	192	6	0.7	4	10	5.8	164	8	5
Cycle 3	186	4	1.3	1	5	6.5	168	10	4
Cycle 4	174	2	0.9	3	5	5.9	164	9	3
Cycle 5	178	3	1.0	2	5	5.7	168	4	0
Phenotypic backcrossing	168		–1.1			5.1	204	13	13
LSD	17		1.2			0.8	33	9	8
Check hybrids									
N03-D8	221					6.1	143	17	6
BIXXIO	261					7.5	135	2	2
39F59	255					7.6	122	35	10

† Rank of plant height or Yield Index relative to the other cycles of selection within the same cross

‡ Yield Index did not include moisture the DC1 × LH74 cross.
§ Least significant difference ($P = 0.05$)

Table 9 Changes in single nucleotide polymorphism (SNP) allele frequency, SNP-based inbreeding coefficient, and linkage disequilibrium (mean r^2 between adjacent SNP markers) for each cycle of genomewide selection in the DC1 × LH74 and DC1 × PHG50 corn crosses.

Cross and cycle	Mean change in SNP allele frequency from previous cycle †	Inbreeding coefficient	Mean r^2 between adjacent SNP markers
DC1 × LH74			
Cycle 0			0.65
Cycle 1	0.11	0.05	0.50
Cycle 2	0.11	0.15	0.43
Cycle 3	0.08	0.22	0.35
Cycle 4	0.08	0.30	0.26
DC1 × PHG50			
Cycle 0			0.67
Cycle 1	0.12	0.11	0.38
Cycle 2	0.12	0.23	0.33
Cycle 3	0.09	0.25	0.34
Cycle 4	0.09	0.36	0.31

†Mean of absolute values of the change in frequency of the SNP alleles from the exotic parent, DC1

Table 10 Testcross performance (with a nondwarf inbred, LH227, as the tester) of different cycles of genomewide selection and of progeny from phenotypic backcrossing in the DC1 × LH74 and DC1 × PHG50 corn crosses.

Cross and cycle	Plant height (cm)	Yield Index	Grain yield (Mg ha^{-1})	Moisture (g kg^{-1})	Root lodging (%)	Stalk lodging (%)
DC1 × LH74						
Cycle 0	243	0.2	5.8	172	13	4
Cycle 1	247	1.8	7.1	179	7	2
Cycle 2	249	1.4	7.3	178	16	6
Cycle 3	258	1.8	7.2	176	8	4
Cycle 4	246	1.1	7.2	191	10	5
Cycle 5	247	0.4	6.8	194	14	7
Phenotypic backcrossing	206	0.6	6.9	145	24	5
LSD†	15	1.3	1.0	37	13	6
DC1 × PHG50						
Cycle 0	213	5.0	5.6	186	12	3
Cycle 1	233	6.1	6.4	187	5	2
Cycle 2	227	5.8	6.2	190	6	2
Cycle 3	225	5.4	6.0	200	8	4
Cycle 4	235	5.3	5.8	211	8	3
Cycle 5	216	4.3	4.8	220	7	3
LSD	16	0.9	0.8	25	7	3

† Least significant difference ($P = 0.05$)

Figure 3 Plot of QTL mapping results for plant height based on F_3 family testcrosses to DC2.

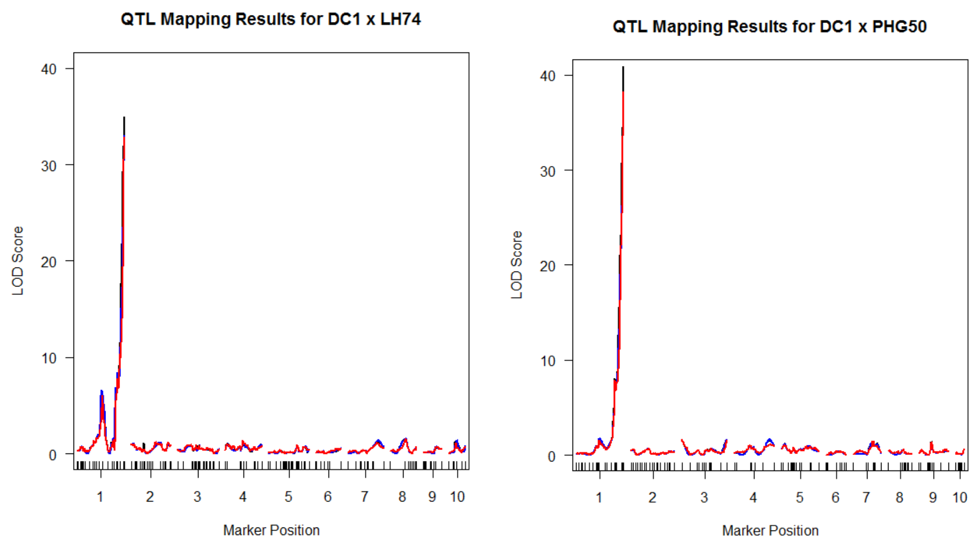
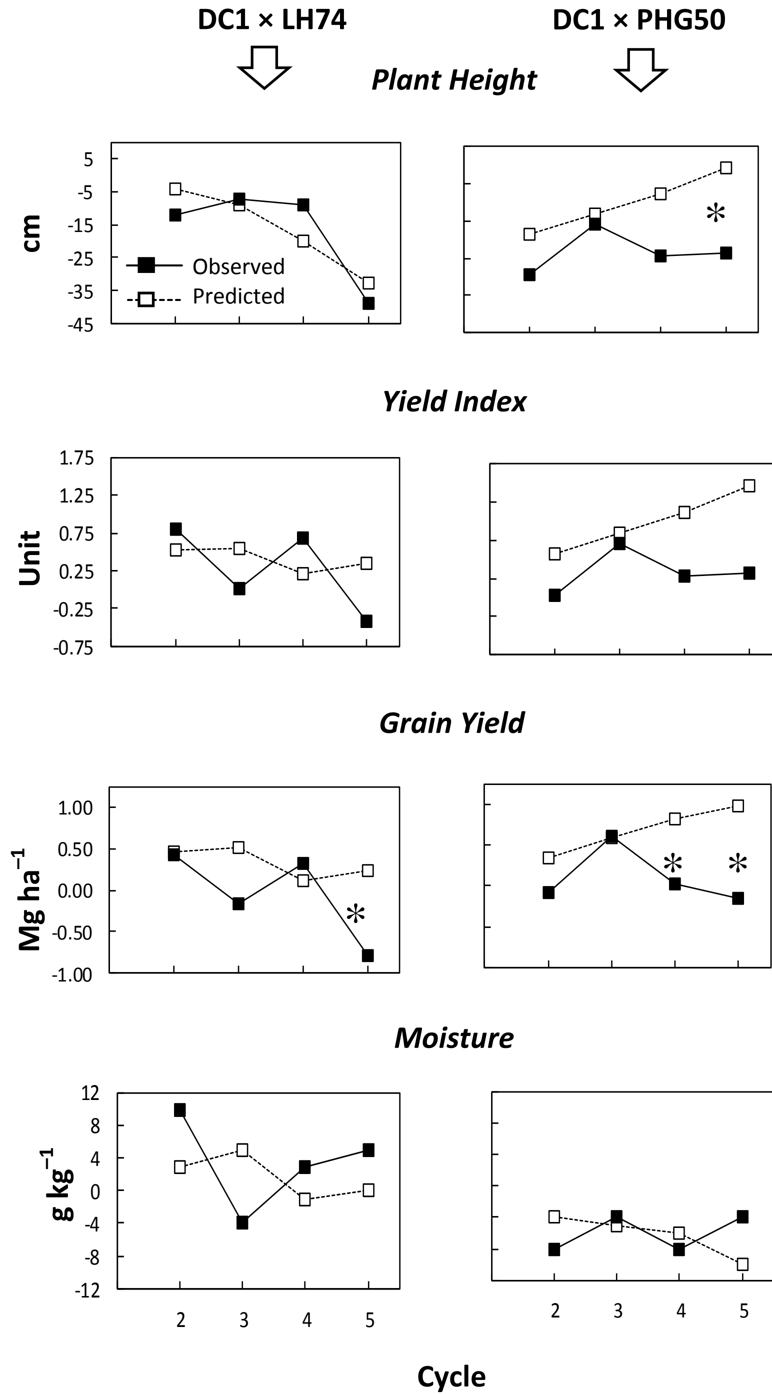


Figure 4 Predicted and observed responses to genomewide selection in the DC1 × LH74 and DC1 × PHG50 corn crosses. A * indicates that the 95% confidence interval on the observed response did not overlap with the predicted response.



Bibliography

Albrecht, B., and J. Dudley. 1987. Evaluation of four maize populations containing different proportions of exotic germplasm. Crop Sci. 27:480-486.

Albrecht, T., V. Wimmer, H. Auinger, M. Erbe, C. Knaak, M. Ouzunova, H. Simianer, and C. Schon. 2011. Genome-based prediction of testcross values in maize. Theor. Appl. Genet. 123:339-350.

Barrett J.C., B. Fry, J. Maller, and M.J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. Jan 15 [PubMed ID: 15297300].

Beavis, W., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor. Appl. Genet. 83:141-145.

Begna, S., R. Hamilton, L. Dwyer, D. Stewart, D. Cloutier, L. Assemat, K. Foroutan-Pour, and D. Smith. 2001. Morphology and yield response to weed pressure by corn hybrids differing in canopy architecture. Eur. J. Agron. 14:293-302.

Bernardo, R. 1991. Retrospective index weights used in multiple trait selection in a maize breeding program. Crop Sci. 31:1174-1179.

Bernardo, R. 1993. Estimation of coefficient of coancestry using molecular markers in maize. Theor. Appl. Genet. 85:1055-1062.

Bernardo, R. 2009. Genomewide selection for rapid introgression of exotic germplasm in maize. Crop Sci. 49:419-425.

Bernardo, R., and J. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47:1082-1090.

Berry, D. P., and J. F. Kearney. 2011. Imputation of genotypes from low- to high-density genotyping platforms and implications for genomic selection. Animal. 5:1162-1169.

Boichard, D., H. Chung, R. Dassonneville, X. David, A. Eggen, S. Fritz, K.J. Gietzen, B.J. Hayes, C.T. Lawley, T.S. Sonstegard, C.P. Van Tassell, P.M. Vanraden, K.A. Viaud-Martinez, and G.R. Wiggans. 2012. Design of a bovine low-density SNP array optimized for imputation. PLoS ONE 7, e34130. doi: 10.1371/journal.pone.0034130.

Broman K.W., H. Wu, Ś. Sen, and G. A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890.

Browning S.R. and B. L. Browning. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. 81:1084-1097.

Calus, M., A. de Roos, and R. Veerkamp. 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178:553-561.

Chia, J.M., C. Song, P.J. Bradbury, D. Costich, N. de Leon, J. Doebley, R.J. Elshire, B. Gaut, L. Geller, and J.C. Glaubitz. 2012. Maize HapMap2 identifies extant variation from a genome in flux. Nat. Genet. 44:803-807.

Cleveland, M.A., J.M. Hickey, and B.P. Kinghorn. 2011. Genotype imputation for the prediction of genomic breeding values in non-genotyped and low-density genotyped individuals. BMC proceedings 5(Suppl 3):S6.

Close, T.J., P.R. Bhat, S. Lonardi, Y. Wu, N. Rostoks, L. Ramsay, A. Druka, N. Stein, J.T. Svensson, and S. Wanamaker. 2009. Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10:582.

Cox, W.J. 1996. Whole-plant physiological and yield responses of maize to plant density. Agron. J. 88:489-496.

Crossa, J., and C. Gardner. 1987. Introgression of an exotic germplasm for improving an adapted maize population. Crop Sci. 27:187-190.

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS One 3:e3395.

Daetwyler, H.D., R. Pong-Wong, B. Villanueva, and J.A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185:1021-1031.

Dassonneville, R., R.F. Brondum, T. Druet, S. Fritz, F. Guillaume, B. Guldbrandtsen, M.S. Lund, V. Ducrocq, and G. Su. 2011. Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. J. Dairy Sci. 94:3679-3686.

Davey, J.W., P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen, and M.L. Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12:499-510.

Dickerson, G.E., C.T. Blunn, and A.G. Chapman. 1954. Evaluation of developing inbred lines of swine. Res. Bull. 551, Missouri Agric. Exp. Stn., Columbia.

Dijak, M., A.M. Modarres, R.I. Hamilton, L.M. Dwyer, D.W. Stewart, D.E. Mather, and D.L. Smith. 1999. Leafy reduced-stature maize hybrids for short-season environments. Crop Sci. 39:1106-1110.

Doerge, R., Z. Zeng and B. Weir. 1994. Statistical issues in the search for genes affecting quantitative traits in populations. p. 15-26. In Statistical issues in the search for genes affecting quantitative traits in populations. Analysis of molecular marker data (supplement). Joint Plant Breed. Symp. Ser., Am. Soc. Hort. Sci., CSSA, Madison, WI.

Dudley, J. 2007. From means to QTL: The illinois long-term selection experiment as a case study in quantitative genetics. Crop Sci. 47:S-20-S-31.

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. The Plant Genome 4:250-255.

Fisher, R.A. 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10:507-521.

Ganal, M.W., G. Durstewitz, A. Polley, A. Bérard, E.S. Buckler, A. Charcosset, J.D. Clarke, E.M. Graner, M. Hansen, and J. Joets. 2011. A large maize (Zea mays L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PloS One 6:e28334.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet.124:323-330.

Goddard, M.E., and B.J. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Rev. Genet. 10:381-391.

Goddard, M.E., and B. Hayes. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. J. Anim. Breed. Genet. 128:409-421.

Goodman, M.M. 1999. Broadening the genetic diversity in maize breeding by use of exotic germplasm. p. 139-148. In J.G. Coors and S. Pandey (eds.). The genetics and exploitation of heterosis in crops. ASA-CSSA-SSSA, Madison, WI.

Gouesnard, B., J. Sanou, A. Panouille, V. Bourion, and A. Boyat. 1996. Evaluation of agronomic traits and analysis of exotic germplasm polymorphism in adapted × exotic maize crosses. Theor. Appl. Genet. 92:368-374.

Grattapaglia, D., and M.D.V. Resende. 2011. Genomic selection in forest tree breeding. Tree Genet. & Genomes 7:241-255.

Grattapaglia, D., C. Plomion, M. Kirst, and R.R. Sederoff. 2009. Genomics of growth traits in forest trees. Curr. Opin. Plant Biol. 12:148-156.

Guo, Z., D.M. Tucker, J. Lu, V. Kishore, and G. Gay. 2012. Evaluation of genome-wide selection efficiency in maize nested association mapping populations. Theor. Appl. Genet.124:261-275.

Hallauer, A.R., and J. Sears. 1972. Integrating exotic germplasm into corn belt maize breeding programs. Crop Sci. 12:203-206.

Hallauer, A.R., and J.B. Miranda, Filho. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames.

Hayes, B., and M. Goddard. 2010. Genome-wide association and genomic selection in animal breeding. Genome 53:876-883.

Hayes, B., P. Bowman, A. Chamberlain, and M. Goddard. 2009. Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433-443.

Hayes, B., P. Bowman, H. Daetwyler, J. Kijas, and J. van der Werf. 2011. Accuracy of genotype imputation in sheep breeds. Anim. Genet. 43:72-80.

Hayes, P.M., B.H. Liu, S.J. Knapp, F. Chen, B. Jones, T. Blake, J. Franckowiak, D. Rasmusson, M. Sorrells, S.E. Ullrich, D. Wesenberg, and A. Kleinhofs. 1993. Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. Theor. Appl. Genet.87:392-401.

Heffner, E.L., J.L. Jannink, and M.E. Sorrells. 2011a. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. The Plant Genome 4:65-75.

Heffner, E.L., J.L. Jannink, H. Iwata, E. Souza, and M.E. Sorrells. 2011b. Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci. 51:2597-2606.

Heffner, E.L., M.E. Sorrells, and J. Jannink. 2009. Genomic selection for crop improvement. Crop Sci. 49:1-12.

Hickey, J.M., J. Crossa, R. Babu, and G. de los Campos. 2012. Factors affecting the accuracy of genotype imputation in populations from several maize breeding programs. Crop Sci. 52:654-663.

Hoffbeck, M., S. Openshaw, J. Geadelmann, R. Peterson, and D. Stuthman. 1995. Backcrossing and intermating in an exotic × adapted cross of maize. Crop Sci. 35:1359-1364.

Holland, J.B. 2007. Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10:156-161.

Iwata, H., and J.L. Jannink. 2010. Marker genotype imputation in a low-marker-density panel with a high-marker-density reference panel: Accuracy evaluation in barley breeding lines. Crop Sci. 50:1269-1278.

Jannink, J.L., H. Iwata, P.R. Bhat, S. Chao, P. Wenzl, and G.J. Muehlbauer. 2009. Marker imputation in barley association studies. The Plant Genome 2:11-22.

Johnson, R. 2004. Marker-assisted selection. Plant Breed. Rev. 24:293-309.

Jumbo, M.D., T. Weldekidan, J.B. Holland, and J.A. Hawk. 2011. Comparison of conventional, modified single seed descent, and doubled haploid breeding methods for maize inbred line development using germplasm enhancement of maize breeding crosses. Crop Sci. 51:1534-1543.

Knapp, S., W. Stroup, and W. Ross. 1985. Exact confidence intervals for heritability on a progeny mean basis. Crop Sci. 25:192-194.

Knott, S.A., and C.S. Haley. 2000. Multitrait least squares for quantitative trait loci detection. Genetics 156:899-911.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Lawrence, C.J., T.E. Seigfried, and V. Brendel. 2005. The maize genetics and genomics database. The community resource for access to diverse maize data. Plant Physiol. 138:55-58.

Lee, M., N. Sharopova, W.D. Beavis, D. Grant, M. Katt, D. Blair, and A. Hallauer. 2002. Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol. Biol. 48:453-461.

Leng, E.R. 1961. Predicted and actual responses during long-term selection for chemical composition in maize. Euphytica 10:368-378.

Lewis, R., and M. Goodman. 2003. Incorporation of tropical maize germplasm into inbred lines derived from temperate× temperate-adapted tropical line crosses: Agronomic and molecular assessment. Theor. Appl. Genet. 107:798-805.

Lewis, M.F., R.E. Lorenzana, H.G. Jung, and R. Bernardo. 2010. Potential for simultaneous improvement of corn grain yield and stover quality for cellulosic ethanol. Crop Sci. 50:516-523.

Lorenz, A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi, H. Iwata, K.P. Smith, M.E. Sorrells and J. Jannink. 2011. Genomic selection in plant breeding: Knowledge and prospects. Adv. Agron. 110:77-123.

Lorenzana, R., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120:151-161.

Lynch, M. 1988. Estimation of relatedness by DNA fingerprinting. Mol. Biol. Evol. 5:584-599.

Marchini, J., B. Howie, S. Myers, G. McVean, and P. Donnelly. 2007. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat. Genet. 39:906-913.

Massman, J.M., H.J.G. Jung, and R. Bernardo. 2012. Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci. doi: 10.2135/cropsci2012.02.0112; published online 21 Aug. 2012.

Mayor, P.J., and R. Bernardo. 2009. Genomewide selection and marker-assisted recurrent selection in doubled haploid versus F2 populations. Crop Sci. 49:1719-1725.

MBS Genetics. 2003. Genetic handbook. 30th ed. Story City, IA.

Meuwissen, T. 2012. The accuracy of genomic selection. 15th European Assoc. Plant Breed. Res. (EUCARPIA) Biometrics in Plant Breed. Section Mtg., 5-7 Sept. 2012, Stuttgart, Germany.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Mikel, M.A., and J.W. Dudley 2006. Evolution of North American dent corn from public to proprietary germplasm. Crop Sci. 46:1193-1205.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z. Zhang, D.E. Costich, and E.S. Buckler. 2009. Association mapping: Critical considerations shift from genotyping to experimental design. The Plant Cell Online 21:2194-2202.

Nelson, P.T., N.D. Coles, J.B. Holland, D.M. Bubeck, S. Smith, and M.M. Goodman. 2008. Molecular characterization of maize inbreds with expired U.S. Plant Variety Protection. Crop Sci. 48:1673-1685.

Nielsen, R., J.S. Paul, A. Albrechtsen, and Y.S. Song. 2011. Genotype and SNP calling from next-generation sequencing data. Nat. Rev. Genet. 12:443-451.

Nothnagel, M., D. Ellinghaus, S. Schreiber, M. Krawczak, and A. Franke. 2009. A comprehensive evaluation of SNP genotype imputation. Hum. Genet. 125:163-171.

Piepho, H.P. 2009. Ridge regression and extensions for genomewide selection in maize. Crop Sci. 49:1165-1176.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS One 7:e32253.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. De Bakker, and M.J. Daly. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559-575.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rawlings, J.O., and D.L. Thompson. 1962. Performance level as criterion for the choice of maize testers. Crop Sci. 2:217-220.

Resende Jr, M., P. Muñoz, M.D.V. Resende, D.J. Garrick, R.L. Fernando, J.M. Davis, E.J. Jokela, T.A. Martin, G.F. Peter, and M. Kirst. 2012. Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda L.). Genetics 190:1503-1510.

Russnogle, J. 1998. Dwarf corn earns tall praise. Corn and Soybean Digest http://cornandsoybeandigest.com/mag/soybean_dwarf_corn_earns/

Sangoi, L. 2001. Understanding plant density effects on maize growth and development: An important issue to maximize grain yield. Ciência Rural 31:159-168.

Schaefer, C.M., C.C. Sheaffer, and R. Bernardo. 2011. Breeding potential of semidwarf corn for grain and forage in the northern U.S. Corn Belt. Crop Sci. 51:1637-1645.

Schaefer, C.M., E. Combs, and R. Bernardo. 2012. Genomewide association mapping and prediction in historical Minnesota maize lines with the Illumina MaizeSNP50 Beadchip. 3rd International Workshop on Next Generation Sequencing Data Analysis and Modern Breeding Approaches, 29-31 Aug. 2012, ICRISAT, Hyderabad, India.

Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. 78:629-644.

Senior, M., E. Chin, M. Lee, J. Smith, and C. Stuber. 1996. Simple sequence repeat markers developed from maize sequences found in the GENBANK database: Map construction. Crop Sci. 36:1676-1683.

Šimić, D., T. Presterl, G. Seitz, and H.H. Geiger. 2003. Comparing methods for integrating exotic germplasm into european forage maize breeding programs. Crop Sci. 43:1952-1959.

Somers, D.J., P. Isaac, and K. Edwards. 2004. A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). Theor. Appl. Genet. 109:1105-1114.

Utz, H.F., A.E. Melchinger, and C.C. Schön. 2000. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839-1849.

Vargas, M., E. Combs, G. Alvarado, G. Atlin, K. Mathews, and J. Crossa. 2012. META: A suite of SAS programs to analyze multienvironment breeding trials. Agron. J. 105:11-19.

Wang, H., B. Woodward, S. Bauck, and R. Rekaya. 2012. Imputation of missing SNP genotypes using low density panels. Livestock Sci. 146:80-83.

Weigel, K.A., G. De los Campos, A.I. Vazquez, G.J.M. Rosa, D. Gianola, and C. P. Van Tassell. 2010. Accuracy of direct genomic values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. J. Dairy Sci. 93:5423-5435.

Zhang, Z., and T. Druet. 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. J. Dairy Sci. 93:5487-5494.

Zhong, S., J.C.M. Dekkers, R.L. Fernando, and J. Jannink. May 2009. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: A barley case study. Genetics 182:355-364.

Ziyomo, C. 2012. Genetic analysis to improve corn grain yield for drought and low nitrogen stress tolerance in a corn-kura clover intercropping system. Ph.D. thesis, Univ. of Minnesota, Saint Paul.

Appendix

Supplemental Table 1: Accuracy of genomewide prediction (r_{MP}) in a biparental maize population with different numbers of markers (N_M), training population sizes (N), and levels of heritability (h^2).

				r_{MP} †			
Trait	N_M	N	Estimated h^2	Estimated h^2	$h^2 = 0.50$	$h^2 =$ 0.30	$h^2 =$ 0.20
Root lodging	1213	48	0.45	0.39		0.33	0.27
Plant height	1213	48	0.74	0.44	0.38	0.32	0.25
Moisture	1213	48	0.85	0.33	0.26	0.20	0.16
Yield	1213	48	0.44	0.23		0.19	0.15
Root lodging	1213	96	0.45	0.50		0.43	0.37
Plant height	1213	96	0.74	0.54	0.47	0.39	0.34
Moisture	1213	96	0.85	0.43	0.35	0.28	0.24
Yield	1213	96	0.44	0.30		0.26	0.21
Root lodging	1213	192	0.45	0.58		0.53	0.46
Plant height	1213	192	0.74	0.61	0.55	0.47	0.42
Moisture	1213	192	0.85	0.51	0.44	0.36	0.30
Yield	1213	192	0.44	0.37		0.33	0.29
Root lodging	256	48	0.45	0.37		0.31	0.25
Plant height	256	48	0.74	0.41	0.35	0.30	0.24
Moisture	256	48	0.85	0.33	0.26	0.21	0.17
Yield	256	48	0.44	0.22		0.18	0.15
Root lodging	256	96	0.45	0.47		0.40	0.35
Plant height	256	96	0.74	0.51	0.44	0.38	0.32
Moisture	256	96	0.85	0.42	0.35	0.28	0.24
Yield	256	96	0.44	0.32		0.25	0.22
Root lodging	256	192	0.45	0.53		0.51	0.43
Plant height	256	192	0.74	0.57	0.51	0.46	0.41
Moisture	256	192	0.85	0.50	0.45	0.37	0.32
Yield	256	192	0.44	0.39		0.33	0.31
Root lodging	512	48	0.45	0.39		0.34	0.27
Plant height	512	48	0.74	0.43	0.38	0.30	0.25
Moisture	512	48	0.85	0.31	0.25	0.20	0.16
Yield	512	48	0.44	0.22		0.18	0.15
Root lodging	512	96	0.45	0.49		0.43	0.37
Plant height	512	96	0.74	0.51	0.46	0.39	0.34

Moisture	512	96	0.85	0.41	0.33	0.27	0.22
Yield	512	96	0.44	0.29		0.25	0.21
Root lodging	512	192	0.45	0.57		0.53	0.45
Plant height	512	192	0.74	0.59	0.53	0.47	0.41
Moisture	512	192	0.85	0.48	0.43	0.36	0.31
Yield	512	192	0.44	0.37		0.32	0.28

† Correlation coefficient, at a given level of h^2, between predicted and observed trait values averaged across 500 runs. LSD = 0.01 (P = 0.05) for each trait.

Supplemental Table 2: Accuracy of genomewide prediction (r_{MP}) in a biparental barley population with different numbers of markers (N_M), training population sizes (N), and levels of heritability (h^2).

Trait	N_M	N	Estimated h^2	r_{MP} †			
				Estimated h^2	$h^2 = 0.50$	$h^2 = 0.30$	$h^2 = 0.20$
Alpha	48	48	0.86	0.68	0.58	0.49	0.41
Extract	48	48	0.88	0.60	0.53	0.46	0.39
Heading date	48	48	0.98	0.37	0.39	0.29	0.25
Lodging	48	48	0.67	0.58	0.53	0.43	0.37
Plant height	48	48	0.96	0.52	0.51	0.43	0.35
Protein	48	48	0.84	0.63	0.51	0.41	0.33
Yield	48	48	0.77	0.39	0.33	0.25	0.20
Alpha	48	72	0.86	0.73	0.64	0.55	0.47
Extract	48	72	0.88	0.66	0.59	0.51	0.44
Head	48	72	0.98	0.55	0.46	0.35	0.29
Plant height	48	72	0.96	0.63	0.58	0.49	0.41
Lodging	48	72	0.67	0.64	0.59	0.50	0.44
Protein	48	72	0.84	0.69	0.59	0.48	0.39
Yield	48	72	0.77	0.46	0.38	0.30	0.25
Alpha	48	96	0.86	0.76	0.68	0.59	0.53
Extract	48	96	0.88	0.69	0.62	0.55	0.49
Heading date	48	96	0.98	0.63	0.52	0.40	0.34
Lodging	48	96	0.67	0.68	0.62	0.54	0.47
Plant height	48	96	0.96	0.69	0.61	0.52	0.46
Protein	48	96	0.84	0.73	0.63	0.53	0.44
Yield	48	96	0.77	0.51	0.42	0.33	0.29
Alpha	100	48	0.86	0.71	0.62	0.53	0.46
Extract	100	48	0.88	0.64	0.57	0.48	0.43
Head	100	48	0.98	0.73	0.57	0.46	0.39
Lodging	100	48	0.67	0.64	0.57	0.47	0.41
Plant height	100	48	0.96	0.77	0.63	0.54	0.47
Protein	100	48	0.84	0.64	0.53	0.43	0.37
Yield	100	48	0.77	0.43	0.37	0.28	0.24
Alpha	100	72	0.86	0.72	0.66	0.60	0.51
Extract	100	72	0.88	0.65	0.60	0.56	0.49
Head	100	72	0.98	0.76	0.65	0.52	0.45
Plant height	100	72	0.96	0.80	0.69	0.60	0.54

Lodging	100	72	0.67	0.68	0.61	0.52	0.46
Protein	100	72	0.84	0.65	0.58	0.49	0.42
Yield	100	72	0.77	0.49	0.43	0.36	0.29
Alpha	100	96	0.86	0.77	0.70	0.63	0.57
Extract	100	96	0.88	0.71	0.65	0.58	0.52
Heading date	100	96	0.98	0.78	0.68	0.59	0.51
Lodging	100	96	0.67	0.72	0.66	0.58	0.51
Plant height	100	96	0.96	0.83	0.73	0.64	0.57
Protein	100	96	0.84	0.71	0.63	0.54	0.46
Yield	100	96	0.77	0.52	0.46	0.39	0.34
Alpha	223	48	0.86	0.72	0.63	0.54	0.47
Extract	223	48	0.88	0.62	0.54	0.48	0.43
Heading date	223	48	0.98	0.67	0.53	0.44	0.36
Lodging	223	48	0.67	0.63	0.57	0.47	0.40
Plant height	223	48	0.96	0.72	0.60	0.52	0.43
Protein	223	48	0.84	0.64	0.53	0.42	0.37
Yield	223	48	0.77	0.39	0.33	0.26	0.22
Alpha	223	72	0.86	0.77	0.68	0.60	0.54
Extract	223	72	0.88	0.67	0.61	0.53	0.47
Head	223	72	0.98	0.80	0.63	0.52	0.44
Plant height	223	72	0.96	0.78	0.66	0.57	0.50
Lodging	223	72	0.67	0.71	0.65	0.55	0.47
Protein	223	72	0.84	0.71	0.61	0.51	0.42
Yield	223	72	0.77	0.47	0.40	0.31	0.26
Alpha	223	96	0.86	0.80	0.73	0.65	0.58
Extract	223	96	0.88	0.70	0.63	0.57	0.51
Heading date	223	96	0.98	0.84	0.68	0.57	0.50
Lodging	223	96	0.67	0.74	0.68	0.59	0.52
Plant height	223	96	0.96	0.82	0.70	0.62	0.55
Protein	223	96	0.84	0.73	0.65	0.56	0.48
Yield	223	96	0.77	0.51	0.44	0.35	0.30

† Correlation coefficient, at a given level of h^2, between predicted and observed trait values averaged across 500 runs. LSD = 0.01 (P = 0.05) for each trait.

Supplemental Table 3: Accuracy of genomewide prediction (r_{MP}) in a mixed barley population with different numbers of markers (N_M), training population sizes (N), and levels of heritability (g^2).

				r_{MP} †			
Trait	N_M	N	Estimated g^2	Estimated g^2	$g^2 =$ 0.50	$g^2 =$ 0.30	$g^2 =$ 0.20
Grain protein	384	72	0.61	0.56	0.53	0.44	0.35
Plant height	384	72	0.72	0.51	0.48	0.44	0.37
Heading date	384	72	0.82	0.49	0.46	0.42	0.39
Grain protein	768	72	0.61	0.58	0.54	0.46	0.36
Plant height	768	72	0.72	0.50	0.48	0.44	0.39
Heading date	768	72	0.82	0.48	0.48	0.43	0.40
Grain protein	1178	72	0.61	0.60	0.55	0.45	0.36
Plant height	1178	72	0.72	0.51	0.49	0.46	0.42
Heading date	1178	72	0.82	0.49	0.48	0.44	0.42

† Correlation coefficient, at a given level of g^2, between predicted and observed trait values averaged across 500 runs. LSD = 0.01 ($P = 0.05$) for each trait.

Supplemental Table 4: Accuracy of genomewide prediction (r_{MP}) in a mixed wheat population with different numbers of markers (N_M), training population sizes (N), and levels of heritability (g^2).

Trait	N_M	N	Estimated g^2	r_{MP} †			
				Estimated g^2	$g^2 = 0.50$	$g^2 = 0.30$	$g^2 = 0.20$
Heading date	384	48	0.95	0.42	0.42	0.38	0.34
Plant height	384	48	0.92	0.40	0.37	0.32	0.28
Maturity	384	48	0.89	0.37	0.36	0.32	0.29
Biomass	384	48	0.38	0.17		0.15	0.13
Yield	384	48	0.68	0.08	0.08	0.07	0.05
Heading date	384	72	0.95	0.41	0.45	0.41	0.37
Plant height	384	72	0.92	0.42	0.40	0.37	0.33
Maturity	384	72	0.89	0.37	0.39	0.36	0.33
Biomass	384	72	0.38	0.20		0.17	0.16
Yield	384	72	0.68	0.09	0.09	0.07	0.06
Heading date	384	96	0.95	0.38	0.46	0.44	0.41
Plant height	384	96	0.92	0.44	0.42	0.39	0.35
Maturity	384	96	0.89	0.38	0.40	0.38	0.35
Biomass	384	96	0.38	0.22		0.19	0.17
Yield	384	96	0.68	0.09	0.10	0.08	0.07
Heading date	576	48	0.95	0.46	0.45	0.40	0.36
Plant height	576	48	0.92	0.48	0.43	0.38	0.34
Maturity	576	48	0.89	0.39	0.37	0.33	0.29
Biomass	576	48	0.38	0.19		0.16	0.15
Yield	576	48	0.68	0.08	0.08	0.06	0.05
Heading date	576	72	0.95	0.45	0.47	0.45	0.40
Plant height	576	72	0.92	0.50	0.47	0.42	0.37
Maturity	576	72	0.89	0.40	0.41	0.38	0.34
Biomass	576	72	0.38	0.21		0.20	0.17
Yield	576	72	0.68	0.09	0.09	0.07	0.06
Heading date	576	96	0.95	0.43	0.48	0.46	0.43
Plant height	576	96	0.92	0.51	0.49	0.45	0.41
Maturity	576	96	0.89	0.40	0.42	0.39	0.37
Biomass	576	96	0.38	0.23		0.20	0.19
Yield	576	96	0.68	0.10	0.09	0.08	0.08

Heading date	731	48	0.95	0.46	0.45	0.40	0.36
Plant height	731	48	0.92	0.49	0.44	0.38	0.33
Maturity	731	48	0.89	0.41	0.38	0.34	0.31
Biomass	731	48	0.38	0.20		0.18	0.17
Yield	731	48	0.68	0.09	0.08	0.06	0.05
Heading date	731	72	0.95	0.46	0.47	0.43	0.40
Plant height	731	72	0.92	0.51	0.46	0.42	0.38
Maturity	731	72	0.89	0.41	0.41	0.38	0.35
Biomass	731	72	0.38	0.22		0.20	0.18
Yield	731	72	0.68	0.09	0.09	0.08	0.06
Heading date	731	96	0.95	0.45	0.49	0.46	0.43
Plant height	731	96	0.92	0.53	0.49	0.45	0.42
Maturity	731	96	0.89	0.42	0.43	0.40	0.38
Biomass	731	96	0.38	0.24		0.23	0.21
Yield	731	96	0.68	0.10	0.10	0.09	0.07

† Correlation coefficient, at a given level of g^2, between predicted and observed trait values averaged across 500 runs. LSD = 0.01 (P = 0.05) for each trait.

Supplemental Table 5: Accuracy of genomewide prediction (r_{MP}) in a simulated maize population with different numbers of markers (N_M), training population sizes (N), and levels of heritability (h^2).

Trait	N_M	N	Estimated h^2	r_{MP} † Estimated h^2	$h^2 = 0.50$	$h^2 = 0.30$	$h^2 = 0.20$
50 QTL	140	48	0.95	0.83	0.67	0.56	0.49
10 QTL	140	48	0.95	0.75	0.58	0.48	0.41
100 QTL	140	48	0.95	0.74	0.58	0.48	0.40
50 QTL	140	96	0.95	0.84	0.75	0.68	0.60
10 QTL	140	96	0.95	0.78	0.69	0.59	0.52
100 QTL	140	96	0.95	0.77	0.68	0.59	0.52
50 QTL	140	192	0.95	0.91	0.83	0.77	0.71
10 QTL	140	192	0.95	0.87	0.78	0.70	0.64
100 QTL	140	192	0.95	0.87	0.78	0.70	0.63
50 QTL	350	48	0.95	0.75	0.66	0.57	0.49
10 QTL	350	48	0.95	0.65	0.56	0.48	0.42
100 QTL	350	48	0.95	0.65	0.55	0.47	0.41
50 QTL	350	96	0.95	0.88	0.77	0.68	0.61
10 QTL	350	96	0.95	0.83	0.70	0.61	0.53
100 QTL	350	96	0.95	0.81	0.69	0.60	0.52
50 QTL	350	192	0.95	0.95	0.84	0.77	0.72
10 QTL	350	192	0.95	0.93	0.80	0.71	0.64
100 QTL	350	192	0.95	0.92	0.78	0.70	0.63
10 QTL	70	48	0.95	0.66	0.55	0.46	0.39
50 QTL	70	48	0.95	0.75	0.64	0.55	0.47
100 QTL	70	48	0.95	0.59	0.53	0.45	0.37
10 QTL	70	96	0.95	0.75	0.67	0.57	0.51
50 QTL	70	96	0.95	0.82	0.74	0.65	0.59
100 QTL	70	96	0.95	0.68	0.63	0.56	0.49
10 QTL	70	192	0.95	0.84	0.75	0.68	0.62
50 QTL	70	192	0.95	0.89	0.80	0.74	0.69
100 QTL	70	192	0.95	0.78	0.71	0.65	0.59

† Correlation coefficient, at a given level of h^2, between predicted and observed trait values averaged across 500 runs. LSD = 0.01 ($P = 0.05$) for each trait.