Published April 28, 2015

RESEARCH

Genomewide Prediction Accuracy within 969 Maize Biparental Populations

Lian Lian, Amy Jacobson, Shengqiang Zhong, and Rex Bernardo*

ABSTRACT

In genomewide selection, the expected correlation between predicted and true genotypic values (r_{MG}) has been previously derived as a function of the training population size (N), heritability (h^2), and effective number of chromosome segments (M_e) affecting the trait. Our objectives were to determine: (i) the mean and variability of r_{MG} in 969 biparental maize (*Zea mays* L.) breeding populations for seven traits, (ii) if r_{MG} can be predicted in advance, and (iii) how N, h^2, and number of markers (N_M) affect r_{MG}. We modified a previous equation for expected r_{MG} to account for linkage disequilibrium (r^2) between a marker and a quantitative trait locus (QTL). Across the 969 populations, the mean and range (in parentheses) of observed r_{MG} was 0.45 (−0.59, 1.03) for grain yield, 0.59 (−0.34, 0.96) for moisture, 0.55 (−0.24, 1.10) for test weight, 0.49 (−0.22, 1.04) for stalk lodging, 0.41 (−0.30, 0.93) for root lodging, 0.47 (−0.45, 0.97) for plant height, and 0.42 (−0.43, 0.94) for ear height. The observed r_{MG} values were centered around the expected r_{MG} when r^2 was accounted for, but the observed r_{MG} had a large spread around the expected r_{MG}. The $r^2(Nh^2)^{1/2}$ had the strongest association with observed r_{MG}. When $r^2(Nh^2)^{1/2}$ exceeded 8, the proportion of r_{MG} equal to or larger than 0.50 reached 90% among all the population–trait combinations. We conclude it is difficult to predict r_{MG} in advance, but that rules of thumb based on $r^2(Nh^2)^{1/2}$ can help achieve a high r_{MG}.

L. Lian, A. Jacobson, and R. Bernardo, Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, 411 Borlaug Hall, 1991 Upper Buford Cir., St. Paul, MN, 55108; and S. Zhong, Monsanto, Ankeny, IA 50021. Received 26 Dec. 2013. *Corresponding author (bernardo@umn.edu).

Abbreviations: DH, doubled haploid; E(r_{MG}), expected prediction accuracy; h^2, heritability; M_e, effective number of chromosome segments; N, population size; N_M, number of markers; QTL, quantitative trait locus; r^2, linkage disequilibrium; r_{MG}, correlation between predicted and true genotypic values; RR-BLUP, ridge regression-best linear unbiased prediction; SNP, single nucleotide polymorphism; V_G, genetic variance; V_R, nongenetic variance.

Maize breeding typically involves crossing two inbred parents, developing inbred progeny from the biparental cross, and selecting the best new inbreds on the basis of testcross performance (Hallauer, 1990). Traditionally, selection for complex traits has been done by field evaluation of testcrosses in multiple environments. Genomewide selection (or genomic selection) (Meuwissen et al., 2001) allows the prediction of genotypic values of individuals for complex traits on the basis of marker information. In genomewide selection, marker effects are estimated from a training population, which has been genotyped and phenotyped. Genotypic values of individuals in a test population, which has been genotyped but not phenotyped, are then predicted from the marker effects estimated from the training population.

Different breeding schemes for genomewide selection have been proposed and studied in maize and in other species (Bernardo, 2009; Bernardo and Yu, 2007; Heffner et al., 2010; Riedelsheimer et al., 2012, 2013; Windhausen et al., 2012). Regardless of the breeding scheme, the r_{MG} must be high enough for

Published in Crop Sci. 54:1514–1522 (2014).
doi: 10.2135/cropsci2013.12.0856
© Crop Science Society of America | 5585 Guilford Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

genomewide selection to be time and cost effective. The expected prediction accuracy [E(r_{MG})] has been previously derived as a function of N, h^2, and M_e affecting the trait (Daetwyler et al., 2008):

$$E(r_{MG}) = \left[Nh^2 / \left(Nh^2 + M_e\right)\right]^{1/2} \quad [1]$$

The M_e pertains to the idealized concept of having independent chromosome segments, with each segment containing a QTL–marker pair and with all the QTL having additive, equal effects (Daetwyler et al., 2008; Goddard, 2009). When the genome is saturated by markers, M_e is generally calculated on the basis of the effective population size and the genome size (Goddard, 2009; Lorenz, 2013; Meuwissen and Goddard, 2010). The M_e can also be calculated by the eigenvalues of the marker correlation matrix, according to the same approach for calculating the effective number of independent tests in association mapping when the markers are highly correlated (Li and Ji, 2005).

Before they commit time and resources to genomewide selection, breeders want to know whether or not r_{MG} will be high enough in a cross. Previous empirical studies on the correspondence between observed and predicted r_{MG} have been few and were limited by the number of populations studied. In five maize biparental crosses, the observed r_{MG} agreed well with the r_{MG} expected from Eq. [1] (Riedelsheimer et al., 2013). In four crosses in maize, barley (*Hordeum vulgare* L.) and *Arabidopsis thaliana* (L.) Heynh, the observed r_{MG} generally agreed with E(r_{MG}) but uncertainty in M_e made the comparisons difficult (Combs and Bernardo, 2013a).

The current study utilizes phenotypic and marker data from 969 biparental maize populations in the Monsanto breeding program. The data in this study therefore represent the genetic backgrounds, traits, h^2, population sizes, and extent of testing in a commercial breeding program. As such, the data can give a realistic indication of the r_{MG} for different traits in maize and of the correspondence between the observed and expected r_{MG} when genomewide selection is routinely practiced on a wide scale. To avoid any confounding effects of a difference in the genetic constitution of the training population and test population (Daetwyler et al., 2008), we analyzed each of the 969 biparental crosses individually. Our objectives were to determine: (i) the mean and variability of r_{MG} in maize biparental populations, (ii) if r_{MG} can be reliably predicted in advance; and (iii) the how r_{MG} is affected by traits, h^2, N, and N_M in biparental populations.

MATERIALS AND METHODS

Phenotypic and Marker Data

The 969 maize populations comprised two maize heterotic groups, with 485 biparental crosses in Group 1 and 484 crosses in Group 2. The number of lines in each cross ranged from 35 to 356 and had a mean of 156. The lines in each cross were derived from the following generations (number of crosses in parentheses): F_2 (707), BC_1 (186), BC_1F_2 (47), doubled haploid (DH) from F_1 (17), and DH from F_2 (12). For the F_2, BC_1, and BC_1F_2 populations, plants grown from 10 to 15 selfed seeds derived from each individual plant were testcrossed to an inbred tester from the opposite heterotic group. For DH populations, each DH line was crossed to an inbred tester.

The populations were evaluated for the following traits: grain yield (Mg ha^{-1}), moisture (g H_2O kg^{-1}), test weight (kg hl^{-1}), stalk lodging (%), root lodging (%), plant height (cm), and ear height (cm). Each experiment was conducted in 2 to 15 locations (usually 6–8 locations) during a single year (2000–2008). Phenotypic data were available as the testcross mean of each line at each location. Not all traits were measured in all locations and in all populations (Table 1). To help ensure that the lodging traits were adequately expressed within a location, only those locations with a mean stalk lodging of at least 5% or a mean root lodging of at least 5% were retained for these two traits, and a population was retained for these two traits only if data were available for more than one location.

The number of polymorphic single nucleotide polymorphism (SNP) markers used to genotype the lines in each population ranged from 31 to 119 and had a mean of 70. The parents of the 969 populations were genotyped with 2911 SNP markers. The lines within each cross were genotyped with 31 to 119 (mean of 70) SNP markers that were polymorphic between the two parents and that were a subset of the 2911 SNP markers used for genotyping all the parents. When the set of markers (31 to 119 SNP loci) used for each population did not have data for some lines, the missing marker data were imputed with fastPHASE (Scheet and Stephens, 2006); imputation with the full set of 2911 parental markers was not done. A segregating SNP locus was disregarded if the parents were monomorphic, if the minor allele frequency was <0.10, or if more than half of the data points were missing. Populations with < 30 markers were removed from analysis.

Heritability

Within each cross, testcross genetic (V_G) and nongenetic (V_R) variance components were calculated for each trait by restricted maximum likelihood via the "lmer" function in the "lme4" package (Bates et al., 2013). Because the data were entry means within each location, the genotype by environment interaction variance and within-location error variance were confounded in V_R. A likelihood ratio test was used to test the significance of the estimates of testcross genetic variance. The p-value from the likelihood ratio test was divided by 2.0 to approximate an F-test of the null hypothesis (Holland et al., 2003). Genetic variance estimates with a p-value > 0.05 were considered not significant and the corresponding population–trait combination was discarded. The data for each biparental cross were not completely balanced because some individuals were not evaluated at a few of the locations in each experiment. As such, the h^2 was estimated on an ad hoc basis as $h^2 = V_G/(V_G + V_R/l)$, where l was the harmonic mean of the number of locations (Holland et al., 2003).

Observed r_{MG}

Genomewide marker effects were obtained by ridge-regression best linear unbiased prediction (RR-BLUP) as described by

Table 1. Mean and range of entry-mean heritability (h^2) and prediction accuracy (r_{MP} and r_{MG}) for different traits in 3371 population–trait combinations in 969 maize biparental crosses.

		h^2		r_{MP}		r_{MG}			
Trait	Populations	Mean	Range	Mean	Range	Mean	Range	50% quantiles	Standard deviation
Yield	840	0.46	(0.17, 0.92)	0.30	(–0.34, 0.89)	0.45a†	(–0.59, 1.03)	(0.32, 0.59)	0.23
Moisture	943	0.66	(0.24, 0.91)	0.48	(–0.18, 0.81)	0.59c	(–0.34, 0.96)	(0.53, 0.71)	0.19
Test weight	894	0.56	(0.18, 0.92)	0.41	(–0.21, 0.78)	0.55b	(–0.24, 1.10)	(0.46, 0.69)	0.20
Stalk lodging	68	0.33	(0.19, 0.67)	0.28	(–0.13, 0.55)	0.49ab	(–0.22, 1.04)	(0.40, 0.64)	0.24
Root lodging	38	0.32	(0.19, 0.53)	0.23	(–0.17, 0.47)	0.4a	(–0.30, 0.93)	(0.21, 0.67)	0.30
Plant height	369	0.39	(0.18, 0.82)	0.29	(–0.27, 0.69)	0.47a	(–0.45, 0.97)	(0.33, 0.62)	0.24
Ear height	219	0.33	(0.17, 0.63)	0.24	(–0.21, 0.59)	0.42a	(–0.43, 0.94)	(0.27, 0.61)	0.25

† r_{MG} values followed by the same letter were not significantly different according to a Tukey HSD test ($P = 0.05$).

Meuwissen et al. (2001) and by Bernardo and Yu (2007). Suppose the phenotypic data comprised N_T records, each record being the mean performance of an individual in one of the N_L locations. The linear model was as $\mathbf{y} = \mathbf{Xb} + \mathbf{Zm} + \mathbf{e}$, where $\mathbf{y}$ was an $N_T \times 1$ vector of phenotypic records, $\mathbf{b}$ was an $N_L \times 1$ vector of fixed effects of locations, $\mathbf{m}$ was an $N_M \times 1$ vector of random effects of the markers, $\mathbf{e}$ was an $N_T \times 1$ vector of residuals, $\mathbf{X}$ was an $N_T \times N_L$ incidence matrix that related $\mathbf{y}$ to $\mathbf{b}$, and $\mathbf{Z}$ was in $N_T \times N_M$ incidence matrix (with values of 1 and −1 for each of the two homozygotes and 0 for the heterozygote) that related $\mathbf{y}$ to $\mathbf{m}$. The variance of marker effects was V_G/N_M (Meuwissen et al., 2001).

The r_{MP} for each trait within each population was estimated by a delete-one method (Kohavi, 1995). Suppose a population had $N_P = 100$ lines that had been phenotyped. The first line was assumed untested and its performance was predicted from genomewide marker effects estimated from RR-BLUP analysis of lines 2 to 100. The second line was then assumed untested and its performance was predicted from RR-BLUP analysis of the remaining $N = N_P - 1$ lines. In the end, the correlation between the predicted genotypic value and the mean phenotypic value of the N_P lines was calculated and was denoted by r_{MP}. The value of r_{MG} was calculated as r_{MP} divided by the square root of h^2 (Dekkers, 2007). For each r_{MP}, we calculated the test statistic $T = [r_{MP}(N_P - 2)^{1/2}]/(1 - r_{MP}^2)]$, which follows a t_{N-2} distribution (Bobko, 2001). Significance tests for r_{MP} (via T) were done through a t test. Given that significance tests for genetic variances had been previously done, we assumed for simplicity that r_{MG} was significantly different from zero if r_{MP} was significantly different from zero.

Expected r_{MG} with Incomplete Linkage Disequilibrium

Equation [1] was derived with the following four assumptions (Daetwyler et al., 2008). First, the marker effects were derived from simple linear regression rather than from RR-BLUP, which we used in this study. Second, each marker–QTL pair was assumed independent of other marker–QTL pairs. Third, the different marker–QTL pairs (which were to be accounted for by M_e) were assumed to have equal variances. Fourth, each marker–QTL pair was assumed in complete linkage disequilibrium. We modified Eq. [1] to retain the first three assumptions by Daetwyler et al. (2008) but to relax the last assumption. By accounting for incomplete linkage disequilibrium between a marker and QTL, we obtained the following $E(r_{MG})$ with $r^2 < 1$ (Appendix 1 in Supplemental Material):

$$E\left(r_{MG\_r2}\right) = r^2\left[Nh^2 / \left(r^2 Nh^2 + M_e\right)\right]^{1/2} \quad [2]$$

In Eq. [2], we used $r^2 = r^2_{MM/2}$, the latter being the mean squared correlation between a marker and QTL when the QTL is assumed to be at the midpoint of the two markers. The midpoint, in turn, is the mean QTL position assuming a uniform distribution between two markers. The r^2_{MM} can be calculated as the square root of the observed squared correlation between two marker genotypes. We found that $r^2_{MM/2} = \sqrt[2]{r^2_{MM}}$ for BC_1, F_2, and DH populations but not for BC_1F_2 populations (Appendix 2 in Supplemental Material). The $E(r_{MG\_r2})$ values were therefore not calculated for the 47 BC_1F_2 populations. For all the following analyses, we used $r^2 = r^2_{MM/2}$. Although the assumption of complete linkage between a marker and a QTL was relaxed, Eq. [2] assumed that the distance between a QTL and a marker is constant for all marker–QTL pairs.

For each population, the M_e values were calculated by eigenvalues of the linkage disequilibrium matrix among the SNP markers in each population (Li and Ji, 2005). This matrix was obtained by computing all pairwise squared correlations for markers on each individual chromosome. Eigenvalue decomposition was done with the "eigen" function in R (R Development Core Team, 2012). Given the known values of N and the estimated values of h^2, r^2, and M_e, the expected r_{MG} was calculated according to both Eq. [1] and [2].

Because of the uncertainty in the proper form of M_e, we also back-calculated M_e by equating the observed r_{MG} with the expected r_{MG} from Eq. [1] and [2] and solving for M_e. This procedure was performed for all population–trait combinations (excluding BC_1F_2) where the observed r_{MG} was between 0.1 and 1.

Factors Affecting r_{MG}

We evaluated the relative importance of the following factors in terms of the variance in r_{MG} explained by each factor: trait, h^2, training population size (N being equal to $N_P - 1$), N_M, heterotic group, and generation type. Given their prominence in Eq. [1] or [2], we evaluated the variance explained by $(Nh^2)^{1/2}$ and $r^2(Nh^2)^{1/2}$. For each trait, we also obtained the correlation coefficient between the observed r_{MG} and the following: $(Nh^2)^{1/2}$, $r^2(Nh^2)^{1/2}$, $E(r_{MG})$ from Eq. [1], and $E(r_{MG\_r2})$ from Eq. [2].

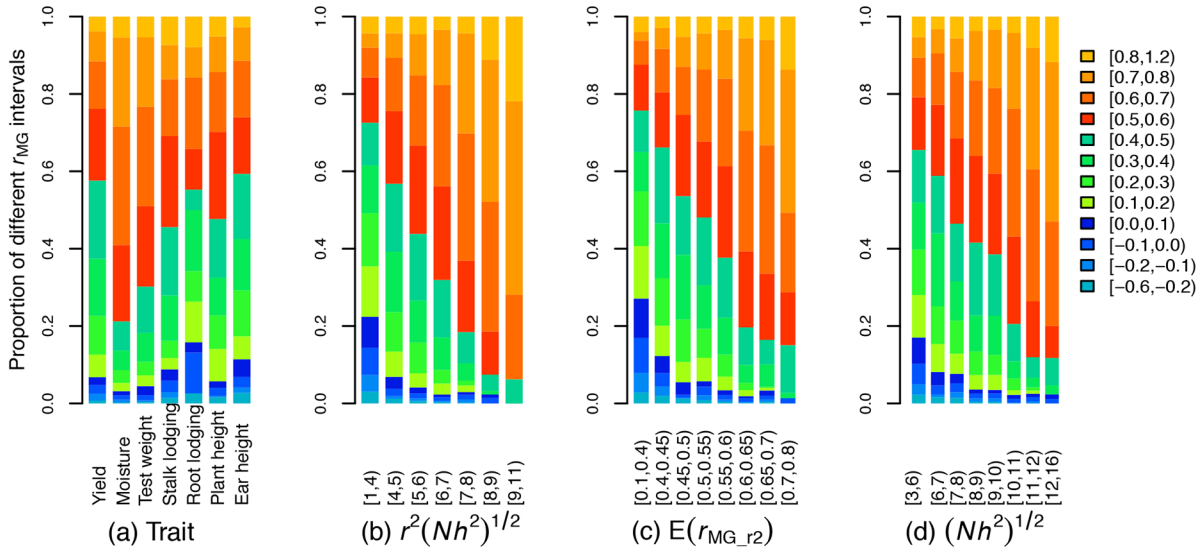

Figure 1. Distribution of observed prediction accuracy (r_{MG}) in biparental maize populations for: (a) different traits for 3371 population–trait combinations, (b) different $r^2(Nh^2)^{1/2}$, (c) different expected r_{MG} from Eq. [2] [$E(r_{MG\_r2})$]; and (d) different $(Nh^2)^{1/2}$. Data in b, c, and d were for 3217 population–trait combinations for which BC_1F_2 populations were excluded. N = training population size, h^2 = heritability on an entry-mean basis, and $r^2 = r^2_{MM/2}$, where $r^2_{MM/2}$ was the mean linkage disequilibrium between a marker and quantitative trait locus (QTL) when the QTL was assumed to be at the midpoint of the two markers.

To further investigate the effect of N on r_{MG}, 1533 population–trait combinations with rounded (to the nearest tens digit) values of N_P = 180 were selected. For each population–trait combination, a random subset of N_P = 61, 91, 121, and 151 lines were obtained as a new population and the r_{MP} within the new population was estimated by a delete-one method as described before, so that N was 60, 90, 120, and 150. For each population–trait combination, only one random sample for the new population was obtained because the results were to be averaged over 1533 populations. The value of h used for calculating r_{MG} (Dekkers, 2007) was that obtained with the largest N_P.

Similarly, to investigate the effect of N_M on r_{MG}, 45 population–trait combinations whose N_M values were larger than 100 were selected. Markers were thinned to 2/3 of their original density by keeping only the first two markers out of every three consecutive markers. Likewise, markers were thinned to 1/3 of their original density by keeping only the first marker out of every three consecutive markers. The r_{MP} within the new population was estimated by a delete-one method as described before.

The effect of h^2 on r_{MG} was further studied for the population–trait combinations with rounded N_P = 180. Heritability values were subdivided into several intervals and the mean r_{MG} in each interval was calculated. For each trait, if the number of populations in an interval was < 20, the mean r_{MG} values for that trait were not calculated for that interval and the data were not shown for that trait in the interval. Because we only had a limited number of observations that met the above criteria for stalk lodging and root lodging, this analysis was not done for these two traits.

Likewise, the effect of $(Nh^2)^{1/2}$ and $r^2(Nh^2)^{1/2}$ on r_{MG} was examined by subdividing $(Nh^2)^{1/2}$ and $r^2(Nh^2)^{1/2}$ into different intervals. For each trait, if the number of populations in an interval was < 20, the mean r_{MG} values for that trait were not calculated for that interval and the data were not shown for that trait in that interval. This analysis was not done for stalk lodging and root lodging and not for BC_1F_2 populations.

To assess their distribution, r_{MG} values were divided into different intervals and the proportion of r_{MG} in each interval was plotted against different traits for all population–trait combinations and plotted against different $(Nh^2)^{1/2}$, different $r^2(Nh^2)^{1/2}$, and different values of $E(r_{MG\_r2})$ for population–trait combinations excluding BC_1F_2.

RESULTS AND DISCUSSION

Mean and Variability of Observed r_{MG}

Out of 3371 population–trait combinations, 2919 (87%) had an r_{MG} that was significantly different from zero (P = 0.05). The mean r_{MG} was 0.52 and the individual r_{MG} values ranged from −0.59 to 1.10 (Table 1). The mean r_{MP} was 0.37 and the individual r_{MP} values ranged from −0.34 to 0.89 (Table 1). The r_{MG} was estimated indirectly as r_{MP}/h (Dekkers, 2007), and the r_{MG} values that exceeded 1.0 were due to sampling variation in both r_{MP} and h^2. Only 11 of the 3371 population–trait combinations (0.3%) had an r_{MG} that was negative and significantly different from zero (P = 0.05).

The mean, range, and standard deviation of r_{MG} values differed among the seven traits. The mean r_{MG} was highest for moisture and lowest for root lodging (Table 1). Conversely, the standard deviation of r_{MG} was smallest for moisture and largest for root lodging. For grain yield, which was the most important trait (Bernardo, 1991), the mean r_{MG} across 840 populations was 0.45 (−0.59, 1.03) and around 40% of the populations had an r_{MG} equal to or larger than 0.50 (Fig. 1a). The middle 50% of r_{MG} values for grain yield ranged from 0.32 to 0.59. For moisture, which was the second most important trait (Bernardo,

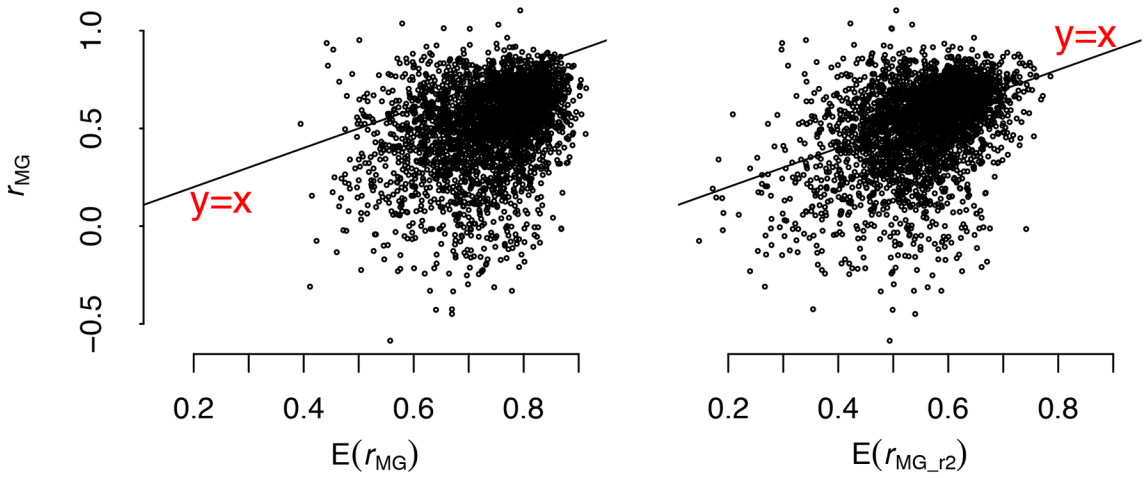

Figure 2. Observed prediction accuracy (r_{MG}) versus expected prediction accuracy from Eq. [1] [E(r_{MG})] and Eq. [2] [E(r_{MG_r2})] for 3217 population–trait combinations (BC_1F_2 populations excluded) in 969 maize biparental crosses.

1991), the mean r_{MG} (range in parentheses) across 943 populations was 0.59 (−0.34, 0.96) and around 80% of the populations had an $r_{MG} \geq 0.50$. The middle 50% of r_{MG} values for moisture ranged from 0.53 to 0.71.

Among the seven traits studied, stalk lodging and root lodging are the two traits that are generally considered by maize breeders as the least consistent in their expression. However, the mean r_{MG} for stalk lodging (0.49) and the mean r_{MG} for root lodging (0.41) were not significantly different from the mean r_{MG} for grain yield (0.45). Overall, the results from this 969-population study indicated that the mean r_{MG} for yield and other agronomic traits in maize biparental crosses is in the 0.40 to 0.60 range.

Expected r_{MG}

There were 3217 population–trait combinations after removing the BC_1F_2 populations, for which r^2 cannot be calculated as the square root of the observed squared correlation between two markers. Whereas the mean r_{MG} across all the 3217 population–trait combinations was 0.52, the mean expected r_{MG} according to Eq. [1] [mean E(r_{MG})] was 0.74. Equation [1] (Daetwyler et al., 2008), which assumes perfect linkage between a marker and QTL, therefore grossly overestimated r_{MG} (by 0.74 − 0.52 = 0.22). The mean expected r_{MG} according to Eq. [2] [mean E(r_{MG_r2})] was 0.56. Equation [2], which accounts for imperfect linkage between a marker and QTL, still overestimated r_{MG}, but the amount of upward bias (0.56 − 0.52 = 0.04) was much less than with Eq. [1] from Daetwyler et al. (2008).

The mean r_{MG} was closer to E(r_{MG_r2}) for some traits than for others. These r_{MG} − E(r_{MG_r2}) deviations were 0.092 for grain yield, 0.004 for moisture, 0.008 for test weight, 0.021 for stalk lodging, 0.102 for root lodging, 0.058 for plant height, and 0.080 for ear height. Across all traits, the spread of r_{MG} about the mean and about E(r_{MG_r2}) was large (Fig. 2).

Pooled across all traits, the correlation between r_{MG} and E(r_{MG_r2}) was 0.40, whereas the correlation between r_{MG} and E(r_{MG}) was 0.33. The correlation between r_{MG} and E(r_{MG_r2}) was significantly different ($P = 0.05$) from the correlation between r_{MG} and E(r_{MG}) for grain yield, moisture, and test weight, but was not significantly different for stalk lodging, root lodging, plant height, and ear height (Table 2). The correlation between r_{MG} and E(r_{MG_r2}) ranged from −0.17 for root lodging to 0.44 for moisture (Table 2). The correlation between r_{MG} and E(r_{MG}) ranged from −0.32 for root lodging to 0.34 for test weight. The correlation between r_{MG} and E(r_{MG_r2}) was lower for grain yield than for moisture and test weight. Likewise, the correlation between r_{MG} and E(r_{MG}) was lower for grain yield than for moisture and test weight.

The usefulness of Eq. [1] and of Eq. [2] for predicting r_{MG} therefore differed among traits. Overall, the results for observed r_{MG} and expected r_{MG} indicated that although

Table 2. Correlation between prediction accuracy (r_{MG}) and different combinations of factors that affect r_{MG} for different traits in 3217 population–trait combinations (BC_1F_2 populations excluded) in 969 maize biparental crosses.

Trait	Populations	$r^2(Nh^2)^{1/2}$	E(r_{MG_r2})	$(Nh^2)^{1/2}$	E(r_{MG})
Yield	799	0.30*b†B‡	0.32*aB	0.21*aA	0.21* bA
Moisture	898	0.45*cB	0.44*bB	0.36*bC	0.32*cdA
Test weight	850	0.43*cB	0.43*bB	0.34*bcA	0.34*dA
Stalk lodging	64	−0.02aA	−0.06cdA	−0.02adA	−0.06aeA
Root lodging	37	−0.04aA	−0.17dA	−0.20dA	−0.32eA
Plant height	358	0.23*abA	0.21*aA	0.22*aA	0.19*abA
Ear height	211	0.20*abA	0.19*acA	0.20*acA	0.18*abcA
All	3217	0.41	0.40	0.36	0.33

* Significant at the 0.05 probability level.

† Within columns (across different traits), correlation values followed by the same lowercase letter were not significantly different according to Fisher z-transformations ($P = 0.05$; Bobko, 2001).

‡ Within rows (across different factor combinations), correlation values followed by the same uppercase letter were not significantly different ($P = 0.05$). Significance tests were done by transforming the correlations to Fisher's z and considering that the correlations were nonindependent because they shared the same common variable r_{MG} (Bobko, 2001).

the mean r_{MG} across many different populations and traits can be predicted fairly well by Eq. [2], the r_{MG} for any given trait in any given population cannot be predicted reliably with either Eq. [1] from Daetwyler et al. (2008) or with Eq. [2], which we derived.

Equation [2] had the following parameters: r^2, N, h^2, and M_e. As shown later, the observed r_{MG} was correlated with $r^2(Nh^2)^{1/2}$, which was the numerator of Eq. [2]. This result suggested that the failure to accurately predict r_{MG} was mainly due to the failure of M_e to adequately mimic the assumptions of equal and additive effects of marker–QTL pairs. We suggest four other reasons for the inability to predict r_{MG}: (i) while we accounted for imperfect linkage disequilibrium between a QTL and a marker in Eq. [2], this imperfect linkage disequilibrium itself led to not all QTL effects being captured and, consequently, to missing h^2 (Makowsky et al., 2011; Manolio et al., 2009) in the model; (ii) the linear additive model did not capture all of the genetic variance due to epistasis not being modeled; (iii) sampling error in the estimates of r_{MP} and h^2, which were used in estimating r_{MG}, contributed to a lower observed correlation between r_{MG} and E(r_{MG_r2}); and (iv) uncertainty remains regarding the proper method to calculate M_e.

Previous studies have calculated M_e as $2N_eL$, where N_e is the effective population size and L is the size of the genome in Morgans (Daetwyler et al., 2008). In a maize biparental population in which all individuals are in theory derived from N_e = 1 F_1 plant, the M_e is approximately 30 when M_e is calculated as $2N_eL$. An M_e of 30 always overestimated r_{MG} in the populations used in this study.

The M_e values we used, which were calculated according to Li and Ji (2005), ranged from 28 to 119 and had a mean of 59 in 3034 population–trait combinations (BC_1F_2 populations excluded and r_{MG} between 0.1 and 1). Because N_M was relatively low, these M_e values did not differ much from the actual N_M, which ranged from 31 to 119 and had a mean of 70. When M_e was back-calculated by equating the observed r_{MG} to E(r_{MG}) from Eq. [1] and solving for M_e, the estimated values of M_e ranged from 242 to 525. These M_e values back-calculated from Eq. [1] were much larger than the M_e calculated according to the Li and Ji (2005) method, larger than the actual N_M, and larger than the value of 30 obtained as $2N_eL$ (Table 2). When M_e was back-calculated by equating the observed r_{MG} to E(r_{MG_r2}) from Eq. [2], the estimated values of M_e ranged from 29 to 110. These M_e values back-calculated from Eq. [2] were of similar magnitude to the M_e calculated according to the Li and Ji (2005) method as well as to N_M (Table 2).

Furthermore, the back-calculated M_e values differed among the traits (Table 3). Among the three traits with the most data, grain yield had the largest back-calculated M_e values whereas moisture and test weight had similar back-calculated M_e values. This result was consistent with the perceived complexity of the traits, with grain yield conceivably being controlled by more QTL compared with moisture and test weight. In contrast, M_e calculated according to the Li and Ji (2005) method or as $2N_eL$ leads to the same M_e across all traits. Overall, these results suggest that other methods need to be developed for calculating M_e.

Table 3. Effective number of chromosome segments (M_e) back-calculated from Eq. [1] and [2] for seven different traits in 3034 population–trait combinations (prediction accuracy between 0.1 and 1; BC_1F_2 populations excluded) from 969 maize biparental populations.

Trait	Populations	Mean N_M†	M_e from Li and Ji (2005)	Mean M_e from Eq. [1]	Mean M_e from Eq. [2]
Yield	742	70	59	468a‡	82c
Moisture	869	70	59	275b	29a
Test weight	811	70	59	283b	38ab
Stalk lodging	57	68	57	242ab	40ac
Root lodging	31	69	57	525ab	110ac
Plant height	338	72	60	434a	86c
Ear height	186	73	61	374ab	72bc

† N_M, number of markers used in genomewide prediction in each biparental population.

‡ Within columns, values followed by the same letter were not significantly different according to a Tukey HSD test (P = 0.05).

Association between r_{MG} and Different Factor Combinations

Whereas the best way to estimate M_e is unclear, N in Eq. [1] and [2] were known and h^2 and r^2 were estimated in this study with well-established procedures. We found that $r^2(Nh^2)^{1/2}$ was most strongly associated with r_{MG}. When only the intercept and a specific factor or a combination of factors was fitted, the variance in r_{MG} explained by the regression model was as follows: $r^2(Nh^2)^{1/2}$, 17.1%; $(Nh^2)^{1/2}$, 12.9%; trait, 8.4%; h^2, 7.8%; N, 4.4%; generation type, 0.4%; heterotic group, 0.0%; and N_M, 0.0%. The importance of $r^2(Nh^2)^{1/2}$ in influencing r_{MG} was in accordance with $r^2(Nh^2)^{1/2}$ being the numerator of Eq. [2]. When $r^2(Nh^2)^{1/2}$ exceeded 8, more than 90% of the r_{MG} values were ≥ 0.50 (Fig. 1b).

The correlation between r_{MG} and $r^2(Nh^2)^{1/2}$ was not significantly different (P = 0.05) from the correlation between r_{MG} and E(r_{MG_r2}) for each trait (Table 2). In other words, r_{MG} was as strongly associated with $r^2(Nh^2)^{1/2}$ as it was with E(r_{MG_r2}). The correlations of the different factors with r_{MG} were higher for grain yield, moisture, and test weight than for stalk lodging, root lodging, plant height, and ear height. Across all the traits, the correlation between r_{MG} and $r^2(Nh^2)^{1/2}$ was 0.41, between r_{MG} and E(r_{MG_r2}) was 0.40, between r_{MG} and $(Nh^2)^{1/2}$ was 0.36, and between r_{MG} and E(r_{MG}) was 0.33. Overall, these results indicated that E(r_{MG_r2}) and $r^2(Nh^2)^{1/2}$ were the best predictors of r_{MG}, and that breeders could manipulate $r^2(Nh^2)^{1/2}$ to get a high r_{MG} (Fig. 1b).

The mean r_{MG} varied among traits when $r^2(Nh^2)^{1/2}$ or $(Nh^2)^{1/2}$ was kept constant (Fig. 3b, c). Grain yield

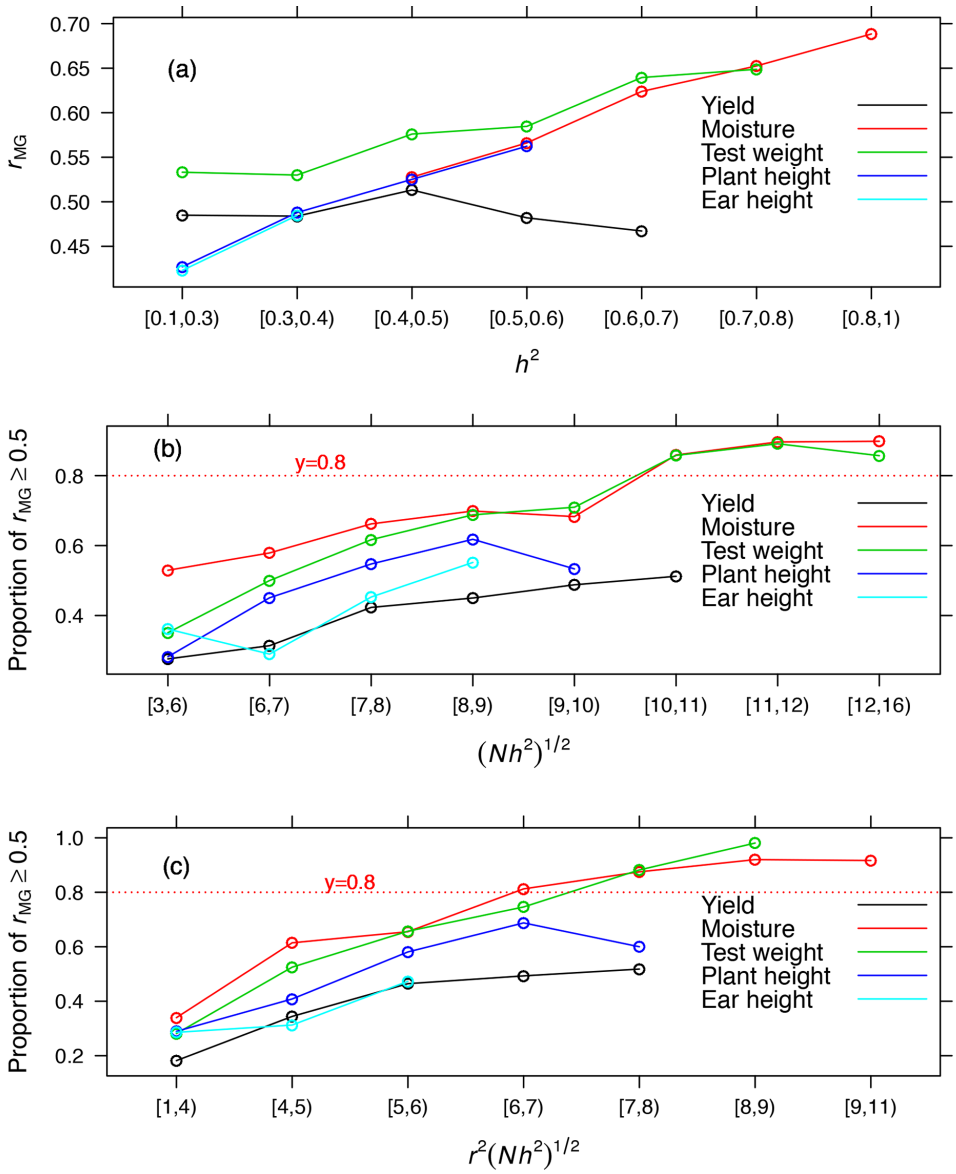

Figure 3. (a) Mean prediction accuracy (r_{MG}) in different intervals of entry-mean heritability (h^2) for five different traits in 1430 population–trait combinations with rounded training population sizes of N_P = 180; (b) proportion of r_{MG} equal to or larger than 0.50 in different intervals of $(Nh^2)^{1/2}$ for five different traits in 3062 population–trait combinations (BC_1F_2 populations excluded); and (c) proportion of r_{MG} equal to or larger than 0.50 in different intervals of $r^2(Nh^2)^{1/2}$ for five different traits in 3073 population–trait combinations (BC_1F_2 populations excluded). N = training population size, $r^2 = r^2_{MM/2}$, where $r^2_{MM/2}$ was the mean linkage disequilibrium between a marker and quantitative trait locus (QTL) when the QTL was assumed to be at the midpoint of the two markers. For all the above plots, population–trait combinations in an interval with <20 data points were excluded.

tended to have a lower r_{MG} compared with moisture, test weight, and plant height at a fixed $r^2(Nh^2)^{1/2}$ or $(Nh^2)^{1/2}$. For grain yield and plant height, the proportion of r_{MG} equal to or larger than 0.50 did not change much beyond $r^2(Nh^2)^{1/2} = 6$ and plateaued at around 50% for grain yield and 60 to 70% for plant height. The proportion of r_{MG} equal to or larger than 0.50 increased to above 80% for moisture and test weight when $r^2(Nh^2)^{1/2}$ exceeded 7.

Association between r_{MG} and Individual Factors

As has been found in previous studies (Combs and Bernardo, 2013a; Crossa et al., 2013; Lorenzana and Bernardo, 2009), increases in the individual factors N, h^2, and N_M generally led to increases in r_{MG}, particularly when it was possible to keep other factors constant. There were 1533 population–trait combinations with the rounded number of lines equal to 180. The mean r_{MG} for different training population sizes was 0.41 with N = 60, 0.46 with N = 90, 0.51 with N = 120, 0.54 with N = 150, and 0.55 with rounded N = 180. These mean r_{MG} values were significantly different from each other (P = 0.05), except for r_{MG} with N = 150 vs. rounded N = 180.

When the rounded N was fixed at 180, r_{MG} increased as h^2 increased for grain yield, moisture, test weight, plant height, and ear height (Fig. 3a). However, the r_{MG} for yield did not change much beyond h^2 = 0.40. In contrast, the r_{MG} for moisture and plant height kept increasing as h^2 increased.

As mentioned above, N_M explained none of the total variance in r_{MG} when other factors were not kept constant. When all other factors were kept constant and the markers were thinned to 2/3 and 1/3 of their original number, the mean r_{MG} was 0.45 when 1/3 of the markers were used (N_M = 36), 0.49 when 2/3 of the markers were used (N_M = 72), and 0.50 when all of the markers were used (N_M = 107). These mean r_{MG} values across different marker densities were not significantly different from each other (P = 0.05). In the above subset of populations, the mean r^2 value between adjacent markers was 0.16 when 1/3 of the markers were used, 0.41 when 2/3 of the markers were used, and 0.52 when all of the markers were used.

A minimum r^2 of 0.20 between adjacent markers has been suggested for genomewide selection (Hayes et al., 2009). The mean r^2 values between adjacent markers was 0.46 across the 969 biparental maize populations used in this study. The r^2 values are typically high in biparental populations because large chromosome segments are passed intact from the inbred parents to the progeny (Smith et al., 2008). The high r^2 between adjacent markers in this study was probably due to having only one meiosis in the development of lines from the F_2 or BC_1 populations, which constituted 92% of the 969 biparental crosses. These high r^2 values suggest that with maize biparental crosses, a fairly small number of SNP markers (N_M = 70 – 120) is largely sufficient for genomewide selection. This agrees with a previous study that found that r_{MG} was at or near maximum when the mean distance between markers was around 25 cM in a DH maize population (Lorenzana and Bernardo, 2009). On the basis of a linkage map of about 1750 cM (Senior et al., 1996), a 25-cM spacing between adjacent markers is equivalent to having 1750/25 = 70 markers. If the linkage map is smaller because of fewer polymorphic markers between the parents, the N_M equivalent to a 25-cM spacing would be even smaller. On the other hand, the N_M needed would be higher for recombinant inbreds, which would have undergone several meiotic events during their development.

Applications in Plant Breeding

Our results for 969 maize populations show that predicting r_{MG} is difficult. The observed r_{MG} values were centered around the expected r_{MG} when recombination between a QTL and marker was accounted for (Eq. [2]), but the spread of the observed r_{MG} around the expected r_{MG} was large. Breeders should also be aware that r_{MG} as well as the ability to predict r_{MG} differ among traits: grain yield tended to have lower r_{MG} and lower predictability of r_{MG} compared with moisture and test weight.

Our results suggest that r_{MG} is best predicted from both $r^2(Nh^2)^{1/2}$ and $E(r_{MG\_r2})$ from Eq. [2]. The correlations with r_{MG} were equal for $r^2(Nh^2)^{1/2}$ and for $E(r_{MG\_r2})$, but the former cannot predict the actual value of r_{MG}. As a rule of thumb, we recommend $r^2(Nh^2)^{1/2}$ to be at least 8. Such a rule of thumb would lead to about 90% of the r_{MG} values exceeding 0.50. When $r^2(Nh^2)^{1/2}$ is between 5 and 6, about 50% or more of the r_{MG} values would exceed 0.50 for most traits. These rules of thumb apply to using a subset of a biparental cross to predict the performance of a remaining, unphenotyped subset of the same biparental cross or for recurrent genomewide selection within the same biparental cross (Combs and Bernardo, 2013b; Massman et al., 2013). Other rules of thumb need to be developed for other types of training populations (e.g., pooled biparental crosses; Jacobson et al., 2014). Also, a marker density of 70 SNP loci seems sufficient for F_2 lines or DH lines developed from an elite biparental cross.

We offer a final thought on predicting r_{MG}: we are unable to precisely predict r_{MG} for each population in the same way that breeders are unable to precisely predict h^2, which measures the effectiveness of phenotypic selection. While we know how increasing the number of replications and environments increases the entry-mean h^2, breeders do not devote time in trying to predict h^2. Instead, a breeder designs a yield trial on the basis of knowledge of how the traits vary in different environments, selects lines often without regard for h^2 in the trial, accepts that the outcome of selection decisions will be poor if h^2 is low, but is confident that selection progress can be made when averaged across different populations. We believe that, likewise, breeders should use information of how different factors affect r_{MG}, design a genomewide selection experiment accordingly, be prepared that the outcome of genomewide-selection decision will be poor if r_{MG} happens to be low in a particular test population, but be confident that routine application of genomewide selection across a breeding program will, on average, lead to positive gains. The results from the 969 maize biparental populations in this study should serve as a useful guide in the design of genomewide selection programs.

Supplemental Information Available

Supplemental information is included with this article.

Acknowledgments

Lian Lian and Amy Jacobson were supported by Ph.D. fellowships funded by Monsanto. We thank David Butruille, Mike Lohuis, and Sam Eathington for allowing us to use the data sets in this study.

References

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2013. Lme4: Linear mixed-effects models using eigen and s4. R package version 1.0–4. R Foundation for Statistical Computing, Vienna, Austria.

Bernardo, R. 1991. Retrospective index weights used in multiple trait selection in a maize breeding program. Crop Sci. 31:1174–1179. doi:10.2135/cropsci1991. 0011183X003100050020x.

Bernardo, R. 2009. Genomewide selection for rapid introgression of exotic germplasm in maize. Crop Sci. 49:419–425. doi:10.2135/cropsci2008.08.0452

Bernardo, R., and J.M. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Sci. 47:1082–1090 doi:10.2135/cropsci2006.11.0690

Bobko, P. 2001. Correlation and regression: Applications for industrial organizational psychology and management. 2nd ed. Sage Publications, Inc., Thousand Oaks, CA.

Combs, E., and R. Bernardo. 2013a. Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. Plant Gen. 6. doi:10.3835/plantgenome2012.11.0030

Combs, E., and R. Bernardo. 2013b. Genomewide selection to introgress semidwarf corn germplasm into U.S. Corn Belt inbreds. Crop Sci. 53:1427–1435. doi:10.2135/cropsci2012.11.0666

Crossa, J., P. Pérez, J. Hickey, J. Burgueño, L. Ornella, J. Cerón-Rojas, et al. 2013. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity 10.1038/hdy.2013.16.

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genomewide approach. PLoS ONE 3:e3395. doi:10.1371/journal.pone.0003395

Dekkers, J. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. J. Anim. Breed. Genet. 124:331–341. doi:10.1111/j.1439-0388.2007.00701.x

Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica (The Hague) 136:245–257. doi:10.1007/s10709–008–9308–0

Hallauer, A. 1990. Methods used in developing maize inbreds. Maydica 35:1–16.

Hayes, B., P. Bowman, A. Chamberlain, and M. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433–443. doi:10.3168/jds.2008–1646.

Heffner, E.L., A.J. Lorenz, J.-L. Jannink, and M.E. Sorrells. 2010. Plant breeding with genomic selection: Gain per unit time and cost. Crop Sci. 50:1681–1690. doi:10.2135/cropsci2009.11.0662

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martínez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9–112. doi:10.1002/9780470650202.ch2.

Jacobson, A., L. Lian, S. Zhong, and R. Bernardo. 2014. General combining ability model for genomewide selection in a biparental cross. Crop Sci. doi:10.2135/cropsci2013.11.0774 (in press).

Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: International joint conference on artificial intelligence, Montreal, Quebec, Canada. 20–25 Aug. American Association for Artificial Intelligence, Montreal. p. 1137–1145.

Li, J., and L. Ji. 2005. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity 95:221–227. doi:10.1038/sj.hdy.6800717

Lorenz, A.J. 2013. Resource allocation for maximizing prediction accuracy and genetic gain of genomic selection in plant breeding: A simulation experiment. G3: Genes| Genomes| Genetics 3:481–491. doi:10.1534/g3.112.004911.

Lorenzana, R.E., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120:151–161. doi:10.1007/s00122-009-1166-3

Makowsky, R., N.M. Pajewski, Y.C. Klimentidis, A.I. Vazquez, C.W. Duarte, D.B. Allison, and G. de los Campos. 2011. Beyond missing heritability: Prediction of complex traits. PLoS Genet. 7. doi:10.1371/journal.pgen.1002051

Manolio, T.A., F.S. Collins, N.J. Cox, D.B. Goldstein, L.A. Hindorff, D.J. Hunter, et al. 2009. Finding the missing heritability of complex diseases. Nature 461:747–753. doi:10.1038/nature08494

Massman, J.M., H.-J.G. Jung, and R. Bernardo. 2013. Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci. 53:58–66. doi:10.2135/cropsci2012.02.0112

Meuwissen, T., and M. Goddard. 2010. Accurate prediction of genetic values for complex traits by whole-genome resequencing. Genetics 185:623–631. doi:10.1534/genetics.110.116590

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

R Development Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Riedelsheimer, C., A. Czedik-Eysenberg, C. Grieder, J. Lisec, F. Technow, R. Sulpice, et al. 2012. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nat. Genet. 44:217–220. doi:10.1038/ng.1033

Riedelsheimer, C., J.B. Endelman, M. Stange, M.E. Sorrells, J.-L. Jannink, and A.E. Melchinger. 2013. Genomic predictability of interconnected biparental maize populations. Genetics 194:493–503. doi:10.1534/genetics.113.150227

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. doi:10.2135/cropsci1996.0011183X003600060043x

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.

Smith, J., T. Hussain, E. Jones, G. Graham, D. Podlich, S. Wall, and M. Williams. 2008. Use of doubled haploids in maize breeding: Implications for intellectual property protection and genetic diversity in hybrid crops. Mol. Breed. 22:51–59. doi:10.1007/s11032-007-9155-1

Windhausen, V.S., G.N. Atlin, J.M. Hickey, J. Crossa, J.-L. Jannink, M.E. Sorrells, et al. 2012. Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. G3: Genes| Genomes| Genetics 2:1427–1436. doi:10.1534/g3.112.003699