Theor Appl Genet (2009) 120:151–161
DOI 10.1007/s00122-009-1166-3

ORIGINAL PAPER

Accuracy of genotypic value predictions for marker-based selection in biparental plant populations

Robenzon E. Lorenzana · Rex Bernardo

Received: 19 February 2009 / Accepted: 27 September 2009 / Published online: 17 October 2009

Abstract The availability of cheap and abundant molecular markers has led to plant-breeding methods that rely on the prediction of genotypic value from marker data, but published information is lacking on the accuracy of genotypic value predictions with empirical data in plants. Our objectives were to (1) determine the accuracy of genotypic value predictions from multiple linear regression (MLR) and genomewide selection via best linear unbiased prediction (BLUP) in biparental plant populations; (2) assess the accuracy of predictions for different numbers of markers (N_M) and progenies (N_P) used in estimation; and (3) determine if an empirical Bayes approach for modeling of the variances of individual markers and of epistatic effects leads to more accurate predictions in empirical data. We divided each of four maize (*Zea mays* L.) datasets, one *Arabidopsis* dataset, and two barley (*Hordeum vulgare* L.) datasets into an estimation set, where marker effects were calculated, and a test set, where genotypic values were predicted based on markers. Predictions were more accurate with BLUP than with MLR. Predictions became more accurate as N_P and N_M increased, until sufficient genome coverage was reached. Modeling marker variances with the empirical Bayes method sometimes led to slightly better predictions, but the accuracy with different variants of the empirical Bayes method was often inconsistent. In nearly all cases, the accuracy with BLUP was not significantly different from the highest accuracy across all methods. Accounting for epistasis in the empirical Bayes procedure led to poorer predictions. We concluded that among the methods considered, the quick and simple BLUP approach is the method of choice for predicting genotypic value in biparental plant populations.

Communicated by M. Cooper.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1166-3) contains supplementary material, which is available to authorized users.

R. E. Lorenzana · R. Bernardo (✉)
Department of Agronomy and Plant Genetics,
University of Minnesota, 411 Borlaug Hall,
1991 Upper Buford Circle,
Saint Paul, MN 55108, USA
e-mail: bernardo@umn.edu

Introduction

The effectiveness of selection depends on the ability to predict the genotypic value from the phenotypic value or from some other selectable criterion. The availability of cheap and abundant molecular markers has led to selection methods that rely on the prediction of genotypic value from marker data (Eathington et al. 2007; Bernardo 2008). Examples of marker-based selection methods for complex traits include marker-assisted recurrent selection (MARS) and selection prior to phenotyping. The MARS approach involves estimating marker effects from phenotypic and marker data in an F_2 population, followed by two or three cycles of selection based only on markers in an off-season nursery or greenhouse where phenotypic data are not meaningful (Edwards and Johnson 1994; Johnson 2004; Eathington et al. 2007). Selection prior to phenotyping involves estimating marker effects from phenotypic and marker databases, followed by marker-based selection in individual F_2 populations (Johnson 2004; Eathington et al. 2007).

Marker effects for complex traits in plants have been typically estimated by least-squares procedures such as multiple linear regression (MLR), and only those markers significantly associated with the traits being improved are

subsequently used in selection (Edwards and Johnson 1994; Johnson 2004). But in a strategy called genomic or genomewide selection, markers are not tested for their significance and all markers are used in selection (Meuwissen et al. 2001). Simulation studies have shown that genotypic value predictions were more accurate with genomewide selection, based on either best linear unbiased prediction (BLUP) or Bayesian methods, than with MLR (Meuwissen et al. 2001). Subsequent simulation studies have shown that selection responses were 18–43% larger with genomewide selection via BLUP than with MARS involving MLR (Bernardo and Yu 2007).

While simulation studies on genomewide selection have involved different numbers of quantitative trait loci (QTL), levels of trait heritability (h^2), and numbers of polymorphic markers (N_M) and progenies (N_P; Meuwissen et al. 2001; Bernardo and Yu 2007), published information is lacking on the accuracy of genomewide predictions with empirical data in plants. The MLR and genomewide selection approaches need to be applied to empirical data so that their usefulness can be compared under the actual genetic models that underlie the expression of quantitative traits in plants. Such comparisons become particularly important because genomewide selection via BLUP involves the unrealistic but convenient assumptions that each of the markers explains an equal proportion of the total genetic variance and that epistasis is absent (Meuwissen et al. 2001). Bayesian approaches for genomewide prediction do not require these two assumptions needed in BLUP (Meuwissen et al. 2001; Xu 2007), but comparisons between Bayesian and BLUP approaches using empirical data in plants are limited.

Our objectives were therefore to (1) determine the accuracy of genotypic value predictions from MLR and genomewide selection via BLUP for different traits in biparental populations of different plant species; (2) assess how N_M and N_P affect the accuracy of predictions in empirical data; and (3) determine if an empirical Bayes approach for modeling of the variances of individual markers and of epistatic effects (Xu 2007) leads to more accurate predictions in empirical data. The accuracy of genotypic value predictions, denoted by r_{MG}, was defined as the correlation between the true genotypic values and the predicted genotypic values based on markers (Meuwissen et al. 2001; Dekkers 2007).

Materials and methods

Phenotypic and genotypic data

Four maize (*Zea mays* L.) datasets (BM-TC1, BM-TC2, LH200BC and Syn-DH), one Arabidopsis (*Arabidopsis thaliana*) dataset (Bay × Sha), and two barley (*Hordeum vulgare* L.) datasets (Steptoe x Morex-DH and Harrington × Morex-DH) were used in this study (Supplementary Table 1). Phenotype data for the BM-TC1 dataset were for 223 intermated B73 × Mo17 recombinant inbreds (Lee et al. 2002) testcrossed to a proprietary Monsanto inbred. The BM-TC1 testcrosses were evaluated at four Minnesota locations in 2007 for grain yield, agronomic traits, and stover quality traits important for cellulosic ethanol production (Lewis 2008). A retrospective index (Yield index) was calculated from grain yield, grain moisture, stalk lodging, and root lodging (Bernardo 1991); a Stover index was calculated as a rank-sum index (Kang 1988) that comprised klason lignin on a cell wall basis (Lignin), glucose concentration (Glucose), and glucose released after pretreatment and enzymatic digestion (Glucose release); and a combined Yield + stover index was calculated as a rank-sum index of the Yield index and Stover index values.

Phenotype data for the BM-TC2 dataset were for 119 intermated B73 × Mo17 recombinant inbreds testcrossed to LH295, an elite Monsanto inbred widely used in the northern U.S. Corn Belt (Lorenzana and Bernardo 2008). The testcrosses were evaluated in an organic production system and conventional production system at two Minnesota locations in 2006. Agronomic traits were measured and a multi-trait selection index for yield was calculated (Table 1).

Genotype data for 1,339 polymorphic markers covering an approximately 6,240 cM linkage map for the intermated B73 × Mo17 population were downloaded from MaizeGDB (Lawrence et al. 2008). Most of the markers were simple sequence repeats (SSR) or restriction fragment length polymorphisms (RFLP).

Phenotype and genotype data for the LH200BC dataset were from Lu et al. (2003). Data were for 349 F_2 progenies derived from LH200 × LH216 and then backcrossed to LH200 after three generations of random mating. The entries were evaluated at five locations in 1999; for this study four locations were used for grain yield and grain moisture, five locations for stalk lodging, and two locations for plant height. Genotype data was available for 160 polymorphic SSR markers covering a 2,580 cM linkage map.

For the Syn-DH dataset, phenotype and genotype data were for testcrosses of 371 doubled haploid lines derived from proprietary Syngenta inbreds (Mayor 2008). The entries were evaluated for grain yield and grain moisture at eight locations and for plant and ear height at two locations in 2006. Genotype data were available for 125 polymorphic single nucleotide polymorphism (SNP) markers covering a 1,490 cM linkage map.

Phenotype and genotype data for the Bay × Sha dataset (Loudet et al. 2002, 2003) were downloaded from the

Table 1 Maximum genotypic value prediction accuracy [max(r_{MG})] for MLR and BLUP for different traits, obtained from the largest N_P in each population

Dataset and trait	BLUP		MLR		$R^c_{BLUP:MLR}$	h^d
	max(r_{MG})	N_M	max(r_{MG})[a]	N^b_M		
BM-TC1 ($N_P = 178$)						
Grain moisture (g kg^{-1})	0.58	1,024	0.36 (0.4)	128 (65)	1.61	0.92
Plant height (cm)	0.70	768	0.52 (0.3)	128 (49)	1.35	0.86
Ear height (cm)	0.49	1,024	0.36 (0.2)	128 (31)	1.36	0.86
Root lodging (%)	0.70	1,024	0.55 (0.2)	256 (69)	1.27	0.82
Stover index	0.65	1,024	0.41 (0.2)	128 (82)	1.59	0.82
Lignin (% cell wall)	0.54	1,024	0.28 (0.2)	256 (80)	1.93	0.81
Yield + stover index	0.70	1,024	0.51 (0.3)	128 (59)	1.37	0.80
Glucose release (%)	0.57	1,024	0.41 (0.2)	128 (34)	1.39	0.76
Yield index	0.73	256	0.59 (0.4)	128 (68)	1.24	0.72
Glucose (g kg^{-1} dry matter)	0.69	512	0.53 (0.2)	128 (37)	1.30	0.71
Grain yield (Mg ha^{-1})	0.61	256	0.45 (0.4)	128 (66)	1.36	0.61
Stalk lodging (%)	0.26	1024	0.15 (0.2)	256 (69)	1.73	0.55
BM-TC2 ($N_P = 119$)						
Grain moisture (g kg^{-1})	0.60	768	0.24 (0.2)	128 (37)	2.50	0.95
Plant height (cm)	0.63	1,024	0.40 (0.2)	128 (37)	1.58	0.90
Ear height (cm)	0.57	768	0.32 (0.2)	128 (37)	1.78	0.82
Yield index	0.54	128	0.35 (0.2)	128 (48)	1.54	0.82
Grain yield (Mg ha^{-1})	0.43	128	0.18 (0.4)	128 (73)	2.39	0.79
Root lodging (%)	0.74	768	0.41 (0.2)	128 (37)	1.80	0.75
LH200BC ($N_P = 279$)						
Grain yield (Mg ha^{-1})	0.72	160	0.65 (0.2)	160 (51)	1.11	0.89
Grain moisture (g kg^{-1})	0.58	160	0.46 (0.4)	160 (83)	1.26	0.88
Plant height (cm)	0.65	128	0.49 (0.2)	128 (35)	1.33	0.69
Stalk lodging (%)	0.50	160	0.41 (0.4)	64 (31)	1.22	0.54
Bay × Sha ($N_P = 332$)						
Flowering time (days)[e]	0.88	69	0.87 (0.4)	69 (33)	1.01	0.93
Free amino acids (nmol mg^{-1} dry matter)[f]	0.91	69	0.86 (0.4)	69 (36)	1.06	0.77
Dry matter at high N (mg plant^{-1})[g]	0.94	69	0.91 (0.4)	69 (37)	1.03	0.67
Dry matter at low N (mg plant^{-1})[h]	0.92	69	0.85 (0.4)	32 (19)	1.12	0.45
Syn-DH ($N_P = 294$ or 297)						
Grain moisture (g kg^{-1})	0.58	64	0.57 (0.3)	64 (25)	1.02 NS	0.94
Grain yield (Mg ha^{-1})	0.38	64	0.32 (0.4)	64 (35)	1.19	0.87
Plant height (cm)	0.58	119	0.50 (0.4)	64 (39)	1.16	0.81
Ear height (cm)	0.46	64	0.47 (0.2)	64 (21)	0.98 NS	0.71
Steptoe x Morex-DH ($N_P = 120$)						
Plant height (cm)	0.86	192	0.85 (0.3)	64 (24)	1.01 NS	0.98
Malt extract (%)	0.76	223	0.66 (0.2)	128 (38)	1.15	0.94
Amylase activity (20 U)	0.88	223	0.83 (0.3)	64 (29)	1.06	0.92
Grain protein content (%)	0.82	128	0.68 (0.2)	128 (45)	1.21	0.92
Grain yield (Mg ha^{-1})	0.62	128	0.54 (0.2)	128 (42)	1.15	0.88
Harrington × Morex-DH ($N_P = 112$)						
Plant height (cm)	0.84	107	0.79 (0.4)	64 (34)	1.06	0.97
Grain protein content (%)	0.73	64	0.74 (0.2)	64 (15)	0.99 NS	0.97
Malt extract (%)	0.70	64	0.68 (0.4)	64 (37)	1.03 NS	0.95

Table 1 continued

Dataset and trait	BLUP		MLR		$R^{c}_{BLUP:MLR}$	h^{d}
	$\max(r_{MG})$	N_M	$\max(r_{MG})^{a}$	N^{b}_M		
Amylase activity (20 U)	0.80	107	0.75 (0.2)	107 (32)	1.07	0.93
Grain yield (Mg ha^{-1})	0.60	107	0.58 (0.2)	64 (17)	1.03 NS	0.91

Traits are arranged in descending order of heritability within each dataset

[a] In parentheses is the significance level that gave the highest r_{MG} in MLR

[b] In parentheses is the mean number of significant markers, at the significance level used, in MLR

[c] Ratio between the maximum accuracy value in BLUP and MLR

[d] Square root of heritability on a progeny-mean basis. All heritabilities were significantly different from zero at $\alpha = 0.05$. Heritabilities in the Bay × Sha dataset were from Loudet et al. (2002, 2003); no significance tests or confidence intervals were reported for these heritabilities

[e] Flowering time under short day regime

[f] Free amino acids content in the shoot in high N condition

[g] Shoot dry matter in high N condition

[h] Shoot dry matter in low N condition

NS Difference in $\max(r_{MG})$ between BLUP and MLR was not significant at $\alpha = 0.05$

Study of the Natural Variation of *A. thaliana* Web page (INRA 2007). Traits selected for analysis included flowering time under short day conditions, free amino acids content in the shoots under high N conditions, shoot dry matter under high N conditions, and shoot dry matter under low N conditions (Table 1). Genotype data for the 415 recombinant inbreds were for 69 polymorphic SSR markers covering a 390 cM linkage map.

For the two barley datasets (Steptoe × Morex-DH and Harrington × Morex-DH), genotype and phenotype data were downloaded from the GrainGenes website (USDA-ARS 2008). Phenotype data for the Steptoe × Morex cross were for 150 doubled haploid lines (Hayes et al. 1993). Grain yield and plant height were measured in 16 environments in 1991–1992, whereas grain protein, malt extract, and alpha amylase activity were measured in nine environments in 1991–1992. Phenotype data for the Harrington × Morex-DH cross were for 140 doubled haploid lines (Marquez-Cedillo et al. 2000). Grain yield and plant height were measured in nine environments in 1995–1996, whereas grain protein, malt extract, and alpha amylase activity were measured in eight environments in 1995–1996. Most of the 223 polymorphic markers covering the 1,250 cM linkage map for the Steptoe x Morex DH population were RFLPs, whereas most of the 107 polymorphic markers covering the 1,000 cM linkage map for the Harrington × Morex DH population were RFLPs and amplified fragment length polymorphism markers.

Estimation of marker effects in MLR

Marker effects were estimated by MLR as described by Bernardo et al. (2006) and Bernardo and Yu (2007). Briefly, multiple regression of the trait values on the number of marker alleles inherited by the individual from the first parental inbred was performed on a per-chromosome basis. Significant markers were identified by backward elimination using relaxed significance levels ($\alpha = 0.2$, 0.3, or 0.4), which were previously found to result in maximum selection response in MARS (Hospital et al. 1997). Marker effects were then estimated as the regression coefficients from a final multiple regression model using all the markers found significant in the per-chromosome analyses.

BLUP of marker breeding values

In genomewide selection via BLUP (referred to simply as BLUP in the rest of this article), marker breeding values were calculated following procedures described by Meuwissen et al. (2001), Muir (2007), and Bernardo and Yu (2007). The phenotypic values for a set of progenies were modeled as

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Xg} + \mathbf{e}$$

where $\mathbf{y}$ was an $N_P \times 1$ vector of phenotypic means of the progenies; $\mathbf{1}$ was an $N_P \times 1$ vector of 1 s; μ was the overall mean; $\mathbf{X}$ was an $N_P \times N_M$ matrix of marker genotype indicators, with elements equal to 1 if the marker locus was homozygous for the allele from the first parent, -1 if the marker locus was homozygous for the allele from the other parent, and 0 if the marker was heterozygous; $\mathbf{g}$ was an $N_M \times 1$ vector of breeding values for the marker alleles from the first parent; and $\mathbf{e}$ was an $N_P \times 1$ vector of residuals. The marker breeding values were obtained by solving the corresponding mixed-model equations (Henderson 1984). Estimates of genetic variance (V_g) and residual variance (V_e) were obtained from an analysis of

variance of phenotypic means across locations. The variance of breeding values at each marker locus was assumed equal to V_g/N_M (Meuwissen et al. 2001).

Estimation of true marker variances and epistatic effects

In the empirical Bayes (e-Bayes) method (Xu 2007), marker main effects and epistatic effects for each pair of markers were calculated based on estimates of true marker variances. Three variations of the e-Bayes method were considered: (1) only marker main effects were estimated and used for predictions (M1); (2) marker main effects and pairwise epistatic effects were estimated, but only the marker main effects were used for predictions (M2); and (3) marker main effects and pairwise epistatic effects were estimated and used for predictions (ME).

Let $\mathbf{z}_l$ denote an $N_P \times 1$ vector of genotype indicators for the lth marker locus. The total number of effects, denoted by q, was equal to N_M if only the main effects were estimated, and equal to $N_M \times (N_M + 1)/2$ if main and pairwise epistatic effects were estimated. Using j to index the jth genetic effect for $j = 1,\ldots,q$, the phenotypic values were modeled as (Xu and Jia 2007)

$$\mathbf{y} = \mathbf{1}\mu + \sum_{j=1}^{q} \mathbf{X}_j g_j + \mathbf{e}$$

where $\mathbf{X}_j$ was equal to $\mathbf{z}_l$ if the jth effect was a main effect or $\mathbf{z}_l \times \mathbf{z}_{l',\, l' > l}$ if the jth effect was an epistatic effect; and g_j was the jth additive or epistatic effect. Variance components for each main effect and epistatic effect were estimated by maximum likelihood. Main and epistatic marker effects given the estimated variance components were then obtained by BLUP (Xu 2007).

The e-Bayes method, like other Bayesian approaches, utilized a prior distribution for the parameters of interest. The prior distribution assigned to the variance of genetic effects was an inverse chi-square distribution with hyperparameters (τ, ω), where τ was the degrees of freedom and ω was the scale parameter (Xu 2007). We used both the model with the optimum hyperparameters (τ, ω) = (-1, 0.0005) and the less optimal [based on simulated data of Xu (2007)] but more computationally efficient hyperparameters (τ, ω) = (-2, 0) for nine selected trait-population combinations. In addition, we used the computationally efficient model with the hyperparameters (τ, ω) = (-2, 0) for 15 other selected trait-population combinations. A SAS PROC IML computer code for implementing the e-Bayes method was obtained from Xu (2007). In the rest of this article, the e-Bayes method with (-1, 0.0005) hyperparameters is referred to as e-Bayes (-1, 0.0005), whereas the e-Bayes method with (-2, 0) hyperparameters is referred to as e-Bayes (-2, 0).

Genotypic value predictions

The predicted genotypic value of an individual, denoted by $\hat{y}_i$, was calculated as (Meuwissen et al. 2001)

$$\hat{y}_i = \hat{\mu} + \mathbf{X}_i\hat{\mathbf{g}}$$

where $\hat{\mu}$ was the estimate of the population mean, $\mathbf{X}_i$ was a row vector of genotype indicators for the ith individual, and $\hat{\mathbf{g}}$ was a column vector of marker effects. When marker effects were calculated by MLR, the elements of $\mathbf{X}_i$ and $\hat{\mathbf{g}}$ included the genotype indicators and marker effects only for the significant markers. When marker main effects were calculated by BLUP or e-Bayes, $\mathbf{X}_i$ was a $1 \times N_M$ vector of genotype indicators and $\hat{\mathbf{g}}$ was a $N_M \times 1$ vector of effects for all the markers. When both marker main effects and epistatic effects were used for predictions, the elements of $\mathbf{X}_i$ were equal to z_{il} for a marker main effect and $z_{il} \times z_{il',\, l' > l}$ for an epistatic effect.

Data analysis

A cross-validation approach similar to the procedures used by Utz et al. (2000), Johnson (2004), and Lee et al. (2008) was used to compare the accuracy of genotypic value predictions from MLR, BLUP, and e-Bayes. For each trait in each population, the data were randomly divided into five subsets. Four subsets were combined to form the estimation set and the remaining subset was used as the test set. Permutation of the five subsets led to five possible sets of estimation and test data per random division of the data.

For each trait, method, and dataset, the correlation between observed phenotypic values and predicted genotypic values (r_{MP}) was calculated in the test set. Estimates of V_g and V_e were obtained from the estimation set. For the Bay × Sha dataset, V_g and V_e were deduced from the published h^2 and the phenotypic variance in the estimation set. The accuracy of genotypic value predictions was expressed as $r_{MG} = r_{MP}/h$ (Dekkers 2007; Lee et al. 2008), where r_{MP} was the correlation between the predicted genotypic values and observed phenotypic values, and h was the square root of heritability on a progeny-mean basis.

To determine the influence of missing marker genotypes (Xu 2007), missing marker genotypes in the maize and barley datasets were first either imputed or set to zero. Missing marker genotypes were imputed using a function in the package R/QTL that used a hidden Markov model to predict the missing marker genotypes given the observed multipoint marker data (Broman et al. 2003). Because the accuracies of the predictions from both approaches were very close (results not shown), the simpler approach of setting the missing marker genotypes to zero was used in all the

analyses. For the BM-TC1 and BM-TC2 datasets, markers with more than 15% missing data were excluded. For the Syn-DH dataset, SNP markers with more than 10% heterozygote genotypes were excluded.

To examine the influence of N_P in the estimation set, random subsets of N_P progenies were selected from each dataset (Supplementary Table 1). To examine the influence of N_M, different numbers of roughly equally spaced markers covering the entire genome were selected. The approximate average distance between adjacent pairs of markers for each N_M was calculated as the length of linkage map (in cM, Supplementary Table 1) divided by the number of marker intervals. The minimum number of markers corresponded to arbitrary marker spacings of approximately 50 cM for the BM-TC1, BM-TC2, and LH200BC datasets, 25 cM for the Syn-DH dataset and the two barley datasets, and 15 cM for the Bay × Sha dataset. Values of N_P used in the estimation set were multiples of 48 whereas values of N_M were multiples of 64 (Table 1; Bernardo and Yu 2007). In MLR, not all marker effects were estimated and no meaningful genotypic value predictions were obtained when the total number of markers found significant in the per-chromosome analyses exceeded N_P.

For the BLUP and MLR comparisons, a total of 100 randomizations (random assignment into five subsets) were performed in each dataset, for a total of 500 total analyses for each method-N_M-N_P combination. The mean of the 500 r_{MP} estimates was obtained. A least significant difference (LSD) at $\alpha = 0.05$ was calculated based on the variance of r_{MP} values across repeats.

For the BLUP and e-Bayes comparisons, fewer randomizations were implemented because of the computing time needed (≥50× longer compared to BLUP). For the BLUP and e-Bayes (−2, 0) comparisons, 20 randomizations for the BM-TC1, BM-TC2, and barley datasets, for a total of 100 analyses, and five randomizations for the other datasets, for a total of 25 analyses, were implemented. For the comparisons of BLUP with the two sets of hyperparameters for the e-Bayes method, five randomizations for a total of 25 analyses were implemented.

All data analyses were done in R (R Development Core Team 2008). The SAS computer code for the e-Bayes method was adapted for R and verified using the datasets and published results in Xu (2007).

Results

Maximum genotypic value prediction accuracy from BLUP and MLR

Genotypic value predictions were generally higher with BLUP than with MLR. The maximum genotypic value prediction accuracy [max(r_{MG})], across values of N_P and N_M, ranged from 0.15 (stalk lodging in the BM-TC1 dataset) to 0.91 (dry matter content at high N in the Bay × Sha dataset) with MLR, and from 0.26 to 0.94 (same traits and datasets) with BLUP (Table 1). With MLR, max(r_{MG}) was equal to or greater than 0.50 for 22 of the 40 trait-dataset combinations. With BLUP, max(r_{MG}) was equal to or greater than 0.50 for 35 of the 40 trait-dataset combinations. The ratio between max(r_{MG}) with BLUP and max(r_{MG}) with MLR was at least 1.05 for 32 of the 40 trait-dataset combinations and was statistically different from 1.0 in 34 out of the 40 trait-dataset combinations.

The ratio between max(r_{MG}) with BLUP and max(r_{MG}) with MLR was generally higher in the intermated populations (BM-TC1, BM-TC2, and LH200BC) than in the non-random-mated populations (Table 1). The max(r_{MG}) generally increased as h^2 increased. A very large N_P, however, compensated for a low h^2 and led to a high max(r_{MG}), e.g., for shoot dry matter at low N in the Bay × Sha dataset. The max(r_{MG}) was close to h for 10 traits in the BM-TC1, BM-TC2, LH200BC, Bay × Sha, and Steptoe × Morex datasets, and max(r_{MG}) was greater than h for three traits in the Bay × Sha dataset.

Genotypic value predictions with different numbers of progenies (N_P)

For all traits the highest r_{MG} for BLUP and MLR were obtained at the highest N_P (Fig. 1, Supplementary Tables 2–8). The decrease in r_{MG} due to a lower N_P was greater with MLR than with BLUP. Consider grain moisture in the BM-TC1 dataset (Supplementary Table 2) and the N_M values that led to max(r_{MG}) with BLUP ($N_M = 1{,}024$) and MLR ($N_M = 128$). The ratio between r_{MG} at $N_P = 96$ and at $N_P = 178$ was 0.81 with BLUP and 0.67 with MLR.

At all values of N_P, BLUP led to r_{MG} values that were greater than or equal to those for MLR across the traits and datasets. The advantage of BLUP over MLR increased as N_P decreased (Supplementary Tables 2, 4–8). Taking grain moisture in the BM-TC1 dataset again as an example, the ratio between the highest r_{MG} with BLUP ($N_M = 1{,}024$) and with MLR ($N_M = 128$) was 1.61 at $N_P = 178$ and 1.96 at $N_P = 96$. Values of N_P needed to obtain r_{MG} values of 0.50 or higher with BLUP for most traits were 144 for the BM-TC1 dataset; 96 for the BM-TC2 dataset; 192 for the Syn-DH and LH200BC datasets; and 48 for the Bay × Sha, Steptoe × Morex-DH, and Harrington x Morex-DH datasets.

Genotypic value predictions with different number of markers (N_M)

The accuracy with MLR generally decreased as N_M increased except when N_P was much larger than N_M

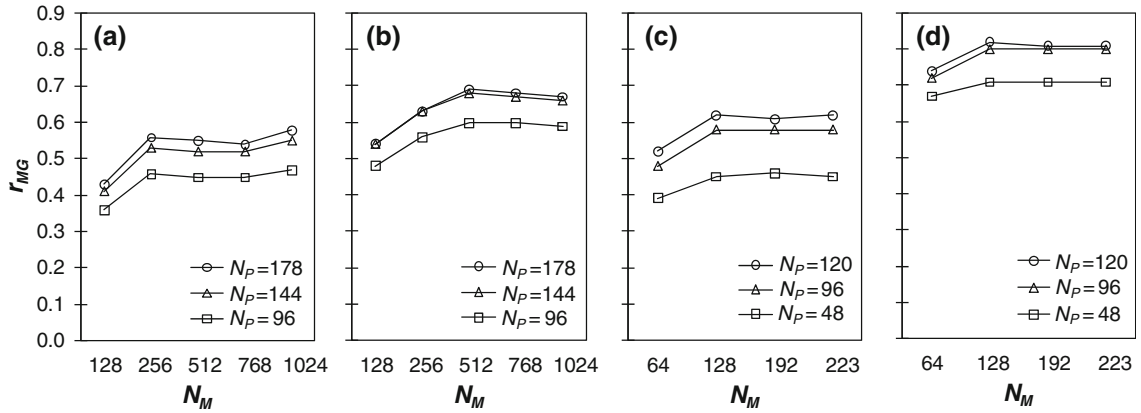

Fig. 1 Genotypic value prediction accuracy (r_{MG}) using BLUP with different number of markers (N_M) and progenies (N_P) in estimating marker effects for selected traits: **a** grain moisture and **b** glucose concentration in corn stover, BM-TC1 dataset; **c** grain yield and **d** grain protein content, Steptoe × Morex-DH dataset

(Supplementary Tables 2–8). For example, increasing N_M from 128 to 256 in the BM-TC1 dataset (Supplementary Table 2), 64–119 in the Syn-DH dataset (Supplementary Table 6), 64–223 in the Steptoe x Morex-DH dataset (Supplementary Table 7), and 64–107 in the Harrington × Morex-DH dataset (Supplementary Table 8) resulted in lower r_{MG} for most of the traits. Exceptions were grain yield and grain moisture in the LH200BC dataset and lignin in the BM-TC1 dataset.

With BLUP, an increase in N_M generally led to increased accuracy (Fig. 1, Supplementary Tables 2–8). For most traits, the highest r_{MG} with BLUP was obtained at the highest N_M, which corresponded to the lowest mean distance between markers (D_M). However, the increase in accuracy became negligible, reached a plateau, or even decreased when a certain level of N_M and corresponding D_M was reached. Consider grain moisture in the BM-TC1 dataset with $N_P = 178$ (Fig. 1a, Supplementary Table 2). The ratio between r_{MG} for $N_M = 256$ ($D_M = 25$ cM) and for $N_M = 128$ ($D_M = 53$ cM) was 1.30, but the r_{MG} values were equal among $N_M = 768$ ($D_M = 8$ cM), $N_M = 512$ ($D_M = 12$ cM), and $N_M = 256$. The ratio between r_{MG} for $N_M = 1{,}024$ ($D_M = 6$ cM) and $N_M = 256$ was only 1.04. Although the highest r_{MG} values with BLUP in the BM-TC1 dataset were mostly obtained at the highest N_M, r_{MG} was at or near maximum with $N_M = 256$ ($D_M = 25$ cM) for grain yield and the Yield index and with $N_M = 512$ or 768 ($D_M = 12$ or 8 cM) for most of the other traits (Supplementary Table 2). For the BM-TC2 dataset, r_{MG} with BLUP was at or near maximum with $N_M = 512$ or 768 ($D_M = 12$ or 8 cM) for most traits, and with $N_M = 128$ ($D_M = 53$ cM) for grain yield and the Yield index (Supplementary Table 3). For the other datasets, r_{MG} for BLUP was at or near maximum with the following N_M values: 160 ($D_M = 17$) in the LH200BC dataset, 64 ($D_M = 6$) in the Bay × Sha dataset, 64 ($D_M = 28$ cM) in the Syn-DH dataset, 128 ($D_M = 10$ cM) in the Steptoe × Morex-DH dataset, and 64 ($D_M = 18$ cM) in the Harrington × Morex-DH dataset (Supplementary Tables 4–8). Although significantly lower than the r_{MG} at the highest N_M, the r_{MG} at $N_M = 128$ ($D_M = 22$) in the LH200BC dataset and $N_M = 32$ ($D_M = 14$) in the Bay × Sha dataset were still high.

Genotypic value predictions when marker variances and epistasis were modeled

The BLUP and e-Bayes comparisons led to contrasting results. Consider the nine trait-population combinations for which both hyperparameters in e-Bayes were used. Of the two e-Bayes variants that did not include epistasis in the model, e-Bayes (−1, 0.0005):M1 led to significantly higher r_{MG} for four trait-population combinations, e-Bayes (−2, 0):M1 led to significantly higher r_{MG} for one trait-population combination, and both procedures led to comparable r_{MG} for four trait-population combinations (Table 2). However, BLUP led to comparable or even higher r_{MG} than either e-Bayes model for eight of the nine trait-population combinations. For the 15 other trait-population combinations where only the (−2, 0) hyperparameters were used for e-Bayes, BLUP provided comparable or greater r_{MG} compared to e-Bayes (−2, 0):M1 (Table 2).

Estimating epistatic effects and including them in the e-Bayes predictions of genotypic value (model M3) consistently led to the poorest predictions (Table 2). For a few traits in the two barley datasets, calculating marker pairwise epistatic effects in addition to the main effects but using only the main effects in predictions (model M2) resulted in improved predictions. For example, r_{MG} was higher with e-Bayes (−2, 0):M2 than with BLUP for two of the eight barley traits (Table 2).

Table 2 Accuracy (r_{MG}) of genotypic value predictions from BLUP and e-Bayes for selected traits at specified N_P and N_M from each dataset

Dataset	Trait	Method[a]							LSD[b]
		BLUP	e-Bayes (−2, 0): M1	e-Bayes (−2, 0): M2	e-Bayes (−2, 0): ME	e-Bayes (−1, 0.0005): M1	e-Bayes (−1, 0.0005): M2	e-Bayes (−1, 0.0005): ME	
BM-TC1 ($N_P = 178$, $N_M = 256$)	Yield index	**0.73**[c]	0.50	0.67	0.56	–	–	–	*0.04*
	Glucose release	**0.48**[d]	**0.52**	**0.49**	0.38	0.39	**0.47**	0.22	*0.09*
	Grain yield	**0.64**	0.38	0.39	0.23	0.51	**0.54**	0.42	*0.11*
BM-TC2 ($N_P = 95$, $N_M = 256$)	Grain moisture	**0.40**	0.23	0.21	0.10	–	–	–	*0.06*
	Plant height	**0.54**	0.36	0.36	0.39	–	–	–	*0.05*
	Yield index	**0.50**	0.29	0.00	-0.06	–	–	–	*0.06*
	Grain yield	**0.40**	0.17	-0.11	-0.01	–	–	–	*0.06*
LH200BC ($N_P = 279$, $N_M = 160$)	Grain yield	**0.72**	**0.70**	0.65	0.51	–	–	–	*0.04*
	Grain moisture	**0.59**	0.53	0.42	0.30	–	–	–	*0.05*
	Stalk lodging	**0.46**	**0.36**	0.21	0.16	–	–	–	*0.13*
Bay × Sha ($N_P = 332$, $N_M = 69$)	Flowering time	**0.88**	**0.88**	**0.86**	0.80	**0.88**	**0.86**	0.85	*0.02*
	Free amino acids	**0.90**	**0.89**	**0.85**	0.71	–	–	–	*0.05*
	Dry matter at high N	**0.93**	**0.93**	0.86	0.75	–	–	–	*0.05*
	Dry matter at low N	**0.91**	**0.83**	0.69	0.67	**0.82**	0.71	0.74	*0.11*
Syn-DH ($N_P = 297$, $N_M = 119$)	Grain moisture	0.55	**0.60**	**0.56**	0.36	–	–	–	*0.04*
	Grain yield	**0.31**	**0.32**	**0.25**	0.07	–	–	–	*0.07*
Steptoe × Morex DH ($N_P = 120$, $N_M = 223$)	Plant height	0.86	0.79	**0.90**	0.84	–	–	–	*0.02*
	Malt extract	**0.76**	0.55	0.66	0.55	–	–	–	*0.04*
	Protein content	**0.83**	0.61	0.70	0.58	**0.79**	0.74	0.12	*0.06*
	Grain yield	**0.64**	0.51	**0.61**	0.40	**0.67**	0.52	−0.02	*0.09*
Harrington × Morex DH ($N_P = 112$, $N_M = 107$)	Plant height	**0.85**	**0.86**	**0.81**	0.72	**0.87**	0.56	0.13	*0.07*
	Protein content	0.66	**0.72**	**0.75**	0.67	**0.78**	0.67	0.63	*0.06*
	Malt extract	**0.69**	**0.68**	0.60	0.53	–	–	–	*0.03*
	Grain yield	**0.61**	0.58	**0.62**	0.48	**0.68**	0.51	0.01	*0.08*

[a] BLUP: best linear unbiased prediction; e-Bayes (−2, 0):M1: marker main effects obtained by e-Bayes (−2, 0) and used for predictions; e-Bayes (−2, 0):M2: marker main effects and pairwise epistatic effects obtained by e-Bayes (−2, 0), but only main effects used for predictions; e-Bayes (−2, 0):ME: marker main effects and pairwise epistatic effects obtained by e-Bayes (−2, 0) and both effects used for predictions; e-Bayes (−1, 0.0005):M1: marker main effects obtained by e-Bayes (−1, 0.0005) and used for predictions; e-Bayes (−1, 0.0005):M2: marker main effects and pairwise epistatic effects obtained by e-Bayes (−1, 0.0005), but only main effects used for predictions; e-Bayes (−1, 0.0005):ME: marker main effects and pairwise epistatic effects obtained by e-Bayes(−1, 0.0005) and both effects used for predictions

[b] Approximate least significant difference at $\alpha = 0.05$

[c] Highest r_{MG} within each trait is set in bold font and underlined

[d] Accuracy values set in bold font are not significantly different from the highest value within each trait

Discussion

Prediction accuracy with BLUP, MLR, and e-Bayes methods

We found four main results in this study. First, genotypic value predictions in empirical data were more accurate when marker effects were calculated by BLUP than by MLR. Second, as expected, genotypic value predictions with BLUP became more accurate as N_P increased and as N_M increased, until sufficient marker coverage of the genome was reached. Third, an empirical Bayes approach that allowed non-equal marker variances led to slightly better predictions, but in many cases BLUP provided comparable or greater prediction accuracy. Finally, including pairwise epistatic interactions in an empirical Bayes approach generally did not improve the accuracy of predictions.

The empirical results for BLUP versus MLR and the influence of N_P and N_M were consistent with previous simulation results (Meuwissen et al. 2001). When N_P was small, the decrease in r_{MG} with both BLUP and MLR was

likely due to the increased colinearity between adjacent markers (Muir 2007) and sampling error (Meuwissen et al. 2001). In addition, the low accuracy of MLR when N_P was small and when N_M was large can be attributed to the poor detection of QTL and overestimation of QTL effects (Beavis 1994; Meuwissen et al. 2001) due to the small degrees of freedom to estimate marker effects (Lande and Thompson 1990; Meuwissen et al. 2001) and colinearity between markers (Whittaker et al. 2000; Gianola et al. 2003). When the marker coverage of the genome was sufficient, N_P became more important in increasing r_{MG}. In simulations, Muir (2007) showed that increased N_M without a corresponding increase in N_P does not increase and can actually decrease r_{MG}. Furthermore, N_P affected the accuracy of not only BLUP and MLR but also of the e-Bayes methods. In particular, the low r_{MG} with e-Bayes (−2, 0) for all traits in the BM-TC2 dataset (Table 2) was likely due to the relatively small population size ($N_P = 95$).

The extent of linkage disequilibrium between markers and QTL determines the number of markers needed for accurate predictions of genotypic value (Meuwissen et al. 2001). Extensive linkage disequilibrium was expected for the non-random-mated biparental plant populations considered in this study (Syn-DH, Bay × Sha, Steptoe × Morex-DH, and Harrington × Morex-DH). Limited recombination events (e.g., one meiosis in doubled haploids) result in large linkage blocks (Smith et al. 2008) and allow the localization of QTL in biparental plant populations only to large intervals (e.g., 10–20 cM) in the genome (Doerge 2002; Holland 2007; Zhu et al. 2008). These results suggest that fewer markers are needed for genomewide prediction of genotypic value in biparental crosses than in open-pollinated populations.

We calculated the effects of individual markers rather than of haplotypes. Hayes et al. (2007) found that calculating marker effects based on haplotypes instead of individual markers increased the accuracy of predictions in cattle (*Bos taurus*). In simulation studies, however, Calus et al. (2008) found that the haplotype approach and individual-marker approach gave comparable r_{MG} when linkage disequilibrium between markers was high. Furthermore, using marker haplotypes has been found most advantageous when N_P was much larger than the population sizes used in this study (Goddard and Hayes 2007; Zhao et al. 2007). In simulations for maize, the use of haplotypes defined in terms of marker intervals led to a slight decrease in the response to several cycles of genomewide selection via BLUP (R. Bernardo, unpublished data, 2008). We therefore speculate that the use of marker haplotypes in this study would not have led to higher r_{MG} values.

Simulation studies (Meuwissen et al. 2001) have shown that modeling true marker variances by Bayesian approaches slightly increased the accuracy of genotypic value predictions compared to BLUP, which involved the convenient but unrealistic assumption of equal marker variances (V_g/N_M; Meuwissen et al. 2001; Bernardo and Yu 2007; Muir 2007). The empirical results in this study generally showed, however, that BLUP provided consistently comparable or even higher accuracy compared to the better of the two sets of e-Bayes hyperparameters. To estimate marker variances in this study, we used an empirical Bayes method, which does not require Markov Chain Monte Carlo (MCMC) samplings (Xu 2007) and is expected to be computationally more efficient than the fully Bayesian methods such as the Stochastic Search Variable Selection method (SSVS, Xu 2007) and BayesB (Meuwissen et al. 2001) which both require MCMC. In a QTL analysis using a doubled haploid barley population with 145 lines and 127 markers, Xu (2007) reported that e-Bayes was more robust than SSVS in that e-Bayes was able to estimate marker effects whereas SSVS led to estimated effects that were all zero. Xu (2007) also argued that the e-Bayes approach is robust with regards to the choice of hyperparameter values.

Inclusion of epistasic effects in marker-assisted selection has been proposed (Lande and Thompson 1990; Dudley 2008), but the empirical results for multiple traits and populations in this study showed no advantage in including pairwise epistatic effects. Results in mice (*Mus musculus*) have likewise showed that including pairwise marker epistatic effects did not improve the accuracy of genomewide predictions (Lee et al. 2008). Results for oil and protein in maize indicated that incorporating epistatic effects improved the accuracy of predictions (Dudley and Johnson 2009). These results for maize, however, involved the identification of markers with statistically significant ($P = 0.05$) main effects and epistatic effects rather than the genomewide approaches we considered for predicting genotypic value.

The e-Bayes procedures that included epistasis (M2 and M3) did not partition the overall genetic variance into the total additive genetic variance and the total epistatic variance. Instead, the variance associated with each main effect and with each pairwise interaction effect was estimated from the data assuming the prior distribution. Including epistatic effects will be useful if two conditions are met: epistasis is important, and epistatic interactions can be modeled accurately. If epistasis is unimportant but the unimportant epistatic interactions can nevertheless be modeled accurately (as being equal to or close to zero), the r_{MG} values should be similar between approaches that do not include epistasis (BLUP, e-Bayes:M1, and e-Bayes:M2) and that include epistasis (e-Bayes:ME). The consistently lower r_{MG} values with e-Bayes:ME in this study indicated the inherent difficulty in genomewide modeling

of epistatic interactions, even if they were present. We speculate that larger population sizes would lead to a better estimation of epistatic variances due to pairs of markers and a better prediction of the corresponding epistatic effects. However, the population sizes we used for the e-Bayes procedures ranged from $N_P = 95$ to 332 and, from a practical standpoint, the use of larger population sizes is unlikely to be routinely feasible in most breeding programs.

Application in breeding programs

The empirical results in this study suggest that among the methods we considered, the quick and simple BLUP approach is the method of choice for genomewide prediction of genotypic value in biparental plant populations. Whereas the two e-Bayes procedures that included epistasis in the predictions were consistently inferior, none of the four other e-Bayes procedures was consistently superior. This lack of consistency makes it difficult to recommend a particular e-Bayes procedure for routine use. On the other hand, in nearly all cases, the r_{MG} with BLUP was not significantly different from the highest accuracy across all methods. Furthermore, the computing time needed for the e-Bayes procedures was about 500 times or more (for the M1 model) to 1,000 times or more (for the M2 and ME models) than that for BLUP.

The efficiency of marker-based selection, relative to phenotypic selection at the same selection intensity, is equal to r_{MG}/h if marker data are assumed to have a heritability of 1.0 (Lande and Thompson 1990; Dekkers 2007). A value of r_{MG} close to h therefore indicates near-equal responses to one cycle of marker-based selection and to one cycle of phenotypic selection. The max(r_{MG}) with BLUP was $\geq$½h for all but two of the 40 trait-dataset combinations in this study (Table 1), indicating that the response to marker-based selection using BLUP would be at least half of the response to phenotypic selection for nearly all traits. Simulation results have likewise indicated that genomewide selection via BLUP accounted for about half of the response to phenotypic selection (Bernardo and Yu 2007).

In temperate regions, only one generation of phenotypic selection can be conducted per year. As indicated in the Introduction, marker-based selection can be conducted in year-round breeding nurseries or greenhouses where phenotypic data are not meaningful but where two to three generations can be grown per year (Edwards and Johnson 1994; Johnson 2004; Eathington et al. 2007). While recombination between markers and QTL may lead to a lower response after the first cycle of genomewide selection (Bernardo and Yu 2007), the max(r_{MG}) values equal to or greater than ½h in this empirical study suggest that the cumulative response from three cycles of genomewide selection via BLUP will approach 1.5 times the gain from one cycle of phenotypic selection. These empirical results therefore indicate that genomewide selection via BLUP would be efficient in terms of genetic gain per year.

When marker coverage of the genome was sufficient and the h^2 for the trait was high, relatively small N_P (i.e., 50 in Arabidopsis to 100–200 in the random-mated maize populations) was needed to reach r_{MG} values of at least 0.50 with BLUP. Larger population sizes would be needed if h^2 is low. The r_{MG} with BLUP was at or near maximum when the mean distance between markers was less than 10 cM in the Arabidopsis dataset, 10–20 cM in the random-mated maize datasets and doubled haploid barley datasets, and around 25 cM in the doubled haploid maize dataset. These values of N_M correspond to <100 markers in the Arabidopsis dataset and doubled haploid maize and barley datasets and about 200–800 markers in the random-mated maize datasets. These numbers of markers can be routinely achieved with high-thoughput SNP genotyping (Jenkins and Gibson 2002; Syvanen 2005) in breeding programs for major crops (Chao et al. 2006; Eathington et al. 2007). Given that phenotyping rather than genotyping is becoming or has become the limiting factor in marker-assisted breeding in major crops (Bernardo 2008), the accuracy of genotypic value predictions in biparental plant populations would likely be influenced mainly by N_P rather than by N_M.

Acknowledgments We thank Monsanto and Syngenta Seeds for access to the BM-TC1 germplasm and Syn-DH data set.

References

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Proceedings of the corn sorghum industrial research conference, 7–8 Dec 1994, Chicago, IL. Am Seed Trade Assoc, Washington DC

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

Bernardo R (2008) Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci 48:1649–1664

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

Bernardo R, Moreau L, Charcosset A (2006) Number and fitness of selected individuals in marker-assisted and phenotypic recurrent selection. Crop Sci 46:1972–1980

Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890

Calus MPL, Meuwissen THE, de Roos APW, Veerkamp RF (2008) Accuracy of genomic selection using different methods to define haplotypes. Genetics 178:553–561

Chao S, Anderson J, Glover K, Smith K (2006) Use of high throughput marker technologies for marker-assisted breeding in wheat and barley. In: Proceedings of plant and animal genome XIV conference, 14–18 Jan 2006, San Diego, CA

Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. J Anim Breed Genet 124:331–341

Doerge RW (2002) Mapping and analysis of quantitative trait loci in experimental populations. Nat Rev Genet 3:43–52

Dudley JW (2008) Epistatic interactions in crosses of Illinois high oil x Illinois low oil and of Illinois high protein x Illinois low protein corn strains. Crop Sci 48:59–68

Dudley JW, Johnson GR (2009) Epistatic models improve prediction of performance in corn. Crop Sci 49:763–770

Eathington SR, Crosbie TM, Edwards MD, Reiter RS, Bull JK (2007) Molecular markers in a commercial breeding program. Crop Sci 47(S3):S154–S163

Edwards M, Johnson L (1994) RFLPs for rapid recurrent selection. p. 33-40. In Analysis of molecular marker data. Joint Plant Breeding Symposia Series. ASA, Madison, WI

Gianola D, Perez-Enciso M, Toro MA (2003) On marker-assisted prediction of genetic value: beyond the ridge. Genetics 163:347–365

Goddard ME, Hayes BJ (2007) Genomic selection. J Anim Breed Genet 124:323–330

Hayes PM, Liu BH, Knapp SJ, Chen F, Jones B, Blake T, Franckowiak J, Rasmusson D, Sorrels M, Ullrich SE, Wesenberg D, Kleinhofs A (1993) Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. Theor Appl Genet 87:392–401

Hayes BJ, Chamberlain AJ, McPartlan H, Macleod I, Sethuraman L, Goddard ME (2007) Accuracy of marker-assisted selection with single markers and marker haplotypes in cattle. Genet Res Camb 89:215–220

Henderson CR (1984) Applications of linear models in animal breeding. Univ of Guelph, ON

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10:156–161

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181–1189

Institut National de la Recherche Agronomique [INRA] (2007) Web Service VNAT, study of the natural variation of *Arabidopsis thaliana*. INRA, Paris, France. Available via http://dbsgap.versailles.inra.fr/vnat/. Accessed 15 January 2009

Jenkins S, Gibson N (2002) High-throughput SNP genotyping. Comp Funct Genom 3:57–66

Johnson R (2004) Marker-assisted selection. Plant Breed Rev 24:293–309

Kang MS (1988) A rank-sum method for selecting high-yielding, stable corn genotypes. Cereal Res Comm 16:113–115

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

Lawrence CJ, Harper LC, Schaeffer ML, Sen TZ, Seigfried TE, Campbell DA (2008) MaizeGDB: The maize model organism database for basic, translational, and applied research. Int J Plant Genom 2008:496957

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

Lee SH, van der Werf JHJ, Hayes BJ, Goddard ME, Visscher PM (2008) Predicting unobserved phenotypes for complex traits from whole-genome SNP data. PLoS Genet 4:e1000231. doi: 10.1371/journal.pgen.1000231

Lewis MF (2008) Potential for simultaneous improvement of maize grain yield and stover quality for cellulosic ethanol. MS thesis, University of Minnesota-Twin Cities, USA

Lorenzana RE, Bernardo R (2008) Genetic correlation between corn performance in organic and conventional production systems. Crop Sci 48:903–910

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F (2002) Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. Theor Appl Genet 104:1173–1184

Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F (2003) Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiol 131:345–358

Lu H, Romero-Severson J, Bernardo R (2003) Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. Theor Appl Genet 107:494–502

Marquez-Cedillo LA, Hayes PM, Jones BL, Kleinhofs A, Legge WG, Rossnagel BG, Sato K, Ullrich E, Wesenberg DM, the North American Barley Genome Mapping Project (2000) QTL analysis of malting quality in barley based on the doubled-haploid progeny of two North American varieties representing different germplasm groups. Theor Appl Genet 101:173–184

Mayor P (2008) Joint use of doubled haploids and molecular markers to increase genetic gain in maize. PhD dissertation, University of Minnesota-Twin Cities, USA

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Muir WM (2007) Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative and genomic parameters. J Anim Breed Genet 124:342–355

Smith JSC, Hussain T, Jones ES, Graham G, Podlich D, Wall S, Williams M (2008) Use of doubled haploids in maize breeding: implications for intellectual property protection and genetic diversity in hybrid crops. Mol Breed 22:51–59

Syvanen AC (2005) Toward genome-wide SNP genotyping. Nat Genet 37:s5–s10

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org

United States Department of Agriculture-Agricultural Research Service [USDA-ARS] (2008) GrainGenes: a database for Triticeae and Avena. Available via http://wheat.pw.usda.gov/GG2/index.shtml. USDA-ARS, Washington, DC. Accessed 4 Oct 2008

Utz HF, Melchinger AE, Schon CC (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

Whittaker JC, Thompson R, Denham MC (2000) Marker-assisted selection using ridge regression. Genet Res Camb 75:249–252

Xu S (2007) An empirical Bayes method for estimating epistatic effects of quantitative trait loci. Biometrics 63:513–521

Xu S, Jia Z (2007) Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics 175:1955–1963

Zhao HH, Fernando RL, Dekkers JCM (2007) Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. Genetics 175:1975–1986

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20