Theor Appl Genet (2011) 123:339–350
DOI 10.1007/s00122-011-1587-7

ORIGINAL PAPER

Genome-based prediction of testcross values in maize

Theresa Albrecht · Valentin Wimmer · Hans-Jürgen Auinger · Malena Erbe · Carsten Knaak · Milena Ouzunova · Henner Simianer · Chris-Carolin Schön

Received: 4 November 2010 / Accepted: 29 March 2011 / Published online: 20 April 2011

Abstract This is the first large-scale experimental study on genome-based prediction of testcross values in an advanced cycle breeding population of maize. The study comprised testcross progenies of 1,380 doubled haploid lines of maize derived from 36 crosses and phenotyped for grain yield and grain dry matter content in seven locations. The lines were genotyped with 1,152 single nucleotide polymorphism markers. Pedigree data were available for three generations. We used best linear unbiased prediction and stratified cross-validation to evaluate the performance of prediction models differing in the modeling of relatedness between inbred lines and in the calculation of genome-based coefficients of similarity. The choice of similarity coefficient did not affect prediction accuracies. Models including genomic information yielded significantly higher prediction accuracies than the model based on pedigree information alone. Average prediction accuracies based on genomic data were high even for a complex trait like grain yield (0.72–0.74) when the cross-validation scheme allowed for a high degree of relatedness between the estimation and the test set. When predictions were performed across distantly related families, prediction accuracies decreased significantly (0.47–0.48). Prediction accuracies decreased with decreasing sample size but were still high when the population size was halved (0.67–0.69). The results from this study are encouraging with respect to genome-based prediction of the genetic value of untested lines in advanced cycle breeding populations and the implementation of genomic selection in the breeding process.

Communicated by J. Snape.

T. Albrecht and V. Wimmer contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1587-7) contains supplementary material, which is available to authorized users.

T. Albrecht · V. Wimmer · H.-J. Auinger · C.-C. Schön (✉)
Plant Breeding, Technische Universität München,
Emil-Ramann-Straße 4, 85354 Freising, Germany
e-mail: chris.schoen@wzw.tum.de

M. Erbe · H. Simianer
Department of Animal Sciences, Georg-August-Universität Göttingen, 37075 Göttingen, Germany

C. Knaak · M. Ouzunova
KWS SAAT AG, 37555 Einbeck, Germany

Introduction

Genomic selection has been widely adopted in animal breeding and is expected to revolutionize breeding methodology (Schaeffer 2006; Jannink et al. 2010). The concept of genomic selection is based on the hypothesis that with a sufficiently high density of genome-wide marker data all of the genetic polymorphisms contributing to trait variation are in high linkage disequilibrium (LD) with random markers segregating in the population under study. Genomic prediction models are developed based on a large training population for which genotypic and phenotypic data are available. The genetic value of untested selection candidates can then be predicted based on their genomic information. The validity of the concept has been shown with computer simulations (Meuwissen et al. 2001) and in experimental studies (Luan et al. 2009; VanRaden et al. 2009). When compared to prediction models using expected values of relatedness based on pedigree data, accuracies of prediction for complex traits like milk yield have increased substantially when the realized

relationships based on molecular marker data are used for estimating breeding values (VanRaden et al. 2009).

In maize breeding, the genetic value of inbred lines is assessed by their testcross performance with testers from the opposite heterotic pool in replicated multi-environment yield trials. Since in a breeding program thousands of lines must be evaluated for their testcross performance each year, models for predictions of the genetic potential based on genotypic data of untested lines are highly desirable. Until recently, prediction of testcross values of untested individuals has not played an important role in plant breeding. Prediction accuracies obtained with expected values of relatedness derived from pedigree data were not sufficient to substitute phenotypic evaluation due to high operative heritabilities achieved with replicated field trials at reasonable costs. However, estimating the realized relationship between individuals from high-density marker data may increase prediction accuracies for testcross performance substantially and thus lead to a paradigm shift in plant breeding as has been observed for cattle breeding. First computer simulations and experimental studies performed with biparental populations in maize and Arabidopsis (Bernardo and Yu 2007; Lorenzana and Bernardo 2009) indicated that genotypic values were predicted with greater accuracy using genome-wide marker data with best linear unbiased prediction (BLUP) compared to marker subset selection with multiple linear regression. However, the fundamental idea of genomic selection is that ancestral functional polymorphisms are tagged by markers and that estimated breeding values are valid for the entire population (Meuwissen et al. 2001). It was further demonstrated by Habier et al. (2010) that prediction accuracy suffers, if the degree of relatedness between the training population and the selection candidates is only weak. Thus, prediction accuracies need to be assessed with experimental material reflecting the genetic structure of advanced cycle breeding populations comprising a large number of crosses and relatively few progenies per family.

To our knowledge, this is the first experimental study in plants reporting accuracies of genomic testcross values estimated at the population level. The dataset in this study consisted of 1,152 genome-wide single nucleotide polymorphism (SNP) markers, pedigree data and phenotypic data on two complex traits, grain yield and grain dry matter content from testcrosses of 1,380 elite maize inbred lines. We used stratified cross-validation to (1) evaluate the prediction accuracy of three approaches to modeling of relatedness between inbred lines, (2) assess the potential of within versus across family prediction, (3) compare the prediction within biparental families to population-wide prediction, and (4) determine the impact of different sample sizes on prediction accuracy.

Materials and methods

Plant material

This study comprised a total of 1,380 doubled haploid (DH) lines of maize (*Zea mays* L.). Thirty-six families were generated from crosses of 29 inbred lines and four single crosses all belonging to the Dent heterotic group. Resulting S_0 plants were used for production of DH lines, which was performed with the in vivo haploid induction technology according to Röber et al. (2005). S_0 plants were pollinated with inducer line RWS and on average 38 DH lines per cross were produced. The smallest DH family comprised 14, the largest 60 lines. For all lines full pedigree information was available up to three generations. The four largest biparental families (BP 1–4), comprising 58–60 DH lines, were also analyzed individually to assess prediction accuracy within individual families.

Field trial design and analysis

All 1,380 DH lines were evaluated as testcrosses with a single-cross Dent tester in 2009 at seven European locations with similar agro-ecological conditions. Two-row plots were machine planted and harvested as grain trials. Data were recorded for grain dry matter yield (GDY, dt/ha) and grain dry matter content (GDC, %). In each of the seven locations trials were performed with one replication and consisted of 15 sets each with 100 entries planted according to a 10×10 lattice design. Each set contained 92 DH lines and four checks replicated twice. Outlying observations were removed from the data set based on extreme deviate standardized residuals according to Grubbs (1950). For each environment, trait values were adjusted for the set effects based on the means of the replicated check varieties.

Genotypic data analysis

Genotyping of the 31 parental inbred lines and their 1,380 DH progenies was performed with 1,152 SNP markers randomly distributed across the genome using the Illumina VeraCode technology. For the majority of the markers, the positions in the maize genome were known from their alignment to the B73 RefGen_v1 sequence (Schnable et al. 2009). The average physical distance between adjacent markers was 2.9 Megabases (Mb). Markers with more than 10% missing values or a minor allele frequency (MAF) < 0.01 were discarded, resulting in 589 and 732 useful SNPs in the population of parental and DH lines, respectively. Three DH lines were discarded from the analysis due to low-quality marker data. Marker genotypes were coded 0 or 2 depending on the number of copies of the

minor allele. Missing marker genotypes were imputed based on family information. If the family from which the DH line was derived did not segregate at the SNP locus, the missing genotype was set to the genotype carried by its siblings. If the SNP marker did segregate in the respective family, the genotype was substituted at random with one of the two possible genotypes at a probability of 0.5. For the population investigated in this study, this method was found to be superior to the approach using flanking markers for imputation (Wimmer et al., unpublished). In addition, it can be applied to markers without known map position.

Linkage disequilibrium estimates were calculated between all pairs of markers over the entire genome. The squared correlation between alleles at two loci was used as a measure of LD (Hill and Robertson 1968):

$$r^2 = \frac{D_{vw}^2}{p_v(1-p_v)p_w(1-p_w)}$$

where $D_{vw} = p_{vw} - p_v p_w$ and p_{vw}, p_v and p_w are the frequencies of the haplotype vw and allele v at one locus and allele w at the other locus. Significance of LD was tested using a χ^2 test according to Foulkes (2009).

Prediction of testcross performance

Analogous to the prediction of breeding values of related individuals in animal breeding, the testcross value of untested DH lines can be predicted if pedigree and/or marker data are available to model the variance–covariance structure between DH lines. Assuming absence of epistasis, the genotypic value of a hybrid individual obtained by crossing random individuals from two parent populations P1 and P2 can be given as

$$y_{i_1 j_2} = \mu_{12} + \alpha_{i_1} + \alpha_{j_2} + \delta_{i_1 j_2} \quad (1)$$

where μ_{12} is the cross population mean, α_{i_1} is the effect of the ith gamete originating from P1, α_{j_2} is the effect of the jth gamete originating from P2, and $\delta_{i_1 j_2}$ is an interaction effect (Schnell 1965; Stuber and Cockerham 1966). Genotypic variances (σ_c^2) and covariances ($\omega_{cc'}$) among cross population relatives can be expressed as

$$\sigma_c^2 = \sigma_{\alpha_1}^2 + \sigma_{\alpha_2}^2 + \sigma_{\delta_{12}}^2 \quad (2)$$

$$\omega_{cc'} = \Phi_1 \sigma_{\alpha_1}^2 + \Phi_2 \sigma_{\alpha_2}^2 + \Phi_1 \Phi_2 \sigma_{\delta_{12}}^2 \quad (3)$$

with $\sigma_{\alpha_1}^2$ and $\sigma_{\alpha_2}^2$ being the variance of general combining ability effects from P1 and P2, $\sigma_{\delta_{12}}^2$ being the interaction variance or variance of specific combining ability effects, and Φ_1 denoting the probability that alleles originating from P1 are identical by descent. In the special case of testcross progenies generated by crossing fully homozygous DH lines originating from the same breeding population (P1) to a common tester from P2 $\sigma_{\alpha_2}^2 = 0$ and $\Phi_2 = 1$ and consequently Eqs. 2 and 3 reduce to

$$\sigma_t^2 = \sigma_{\alpha_1}^2 + \sigma_{\delta_{12}}^2 \quad (4)$$

$$\omega_{tt'} = \Phi_1 \left(\sigma_{\alpha_1}^2 + \sigma_{\delta_{12}}^2 \right) = \Phi_1 \sigma_t^2 \quad (5)$$

with σ_t^2 being the variance of unrelated testcross progenies from P1. When only one common tester is used, the variance components due to general and specific combining ability effects cannot be estimated independently. The magnitude of $\sigma_{\delta_{12}}^2$ depends on the type of tester and equals zero only if the tester represents the entire gametic array of population P2.

To predict testcross values of DH lines, we used mixed effects models and best linear unbiased prediction (Henderson 1984). In all models, the vector of phenotypes **y** comprised the adjusted mean testcross performance averaged across the seven locations of the $N = 1{,}377$ DH lines with high-quality genotypic data. Models differed with respect to modeling the variance–covariance structure of random testcross effects.

Model 1

In Model 1, the probability that two DH lines carry alleles identical by descent was computed based on pedigree information. The vector of adjusted testcross means **y** was modeled as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{t} + \mathbf{e}$$

where $\boldsymbol{\beta}$ is a vector of fixed effects and **X** is a design matrix assigning fixed effects to the phenotypes. Fixed effects include only the population mean and therefore **X** is a vector of 1s. The vector **t** is a vector of testcross effects following a normal distribution with $\mathbf{t} \sim N(0, \mathbf{K}\sigma_t^2)$, where **K** is the $N \times N$ kinship matrix of DH lines calculated from three generations of pedigree data and σ_t^2 is the testcross variance as defined in Eq. 4 pertaining to Model 1. **Z** is a design matrix assigning the genetic testcross values to the phenotypes. The **K** matrix was constructed following Bernardo (2002). When the parents of S_0 plants from which DH lines were derived were not fully homozygous, it was assumed that only one gamete per S_0 plant was sampled and consequently only one DH line was produced from each S_0 plant. Residual effects in the vector **e** are assumed to be independent and follow a normal distribution with $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2)$, where **I** is an identity matrix and σ^2 is the residual variance.

Model 2

In Model 2, the variance–covariance structure of testcross effects was modeled based on marker information. For

calculating the realized kinship matrix between the DH lines we adapted the method proposed by Habier et al. (2007) and VanRaden (2008) to fully homozygous inbred lines. The vector of adjusted testcross means is then modeled as

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$$

with $\mathbf{u}$ being the vector of testcross values following a normal distribution $\mathbf{u} \sim N(0, \mathbf{U}\sigma_u^2)$, where σ_u^2 is the testcross variance pertaining to Model 2. Vectors $\boldsymbol{\beta}$ and $\mathbf{e}$ and design matrices $\mathbf{X}$ and $\mathbf{Z}$ are defined as in Model 1. The realized kinship matrix $\mathbf{U}$ in Model 2 was computed as

$$\mathbf{U} = \frac{(\mathbf{W} - \mathbf{P})(\mathbf{W} - \mathbf{P})'}{8\sum_{m=1}^{M} p_m(1 - p_m)}$$

where $\mathbf{W}$ is an $N \times M$ design matrix assigning $M = 732$ SNP marker genotypes coded 0 or 2 to N phenotypes, $\mathbf{P}$ comprises M column vectors containing the expected genotype score at marker locus m which is a function of the allele frequency, i.e. $2p_m$ for homozygous inbred lines with p_m being the allele frequency of the minor allele at marker locus m. Following Habier et al. (2007) and considering that with fully homozygous inbred lines the variance at individual marker loci equals $4\,p_m(1 - p_m)$ division by $8\sum_{m=1}^{M} p_m(1 - p_m)$ scales the matrix $\mathbf{U}$ to be analogous to the kinship matrix between DH lines.

Model 3

Model 3 combines the realized and the expected kinship in one model. The testcross variance is decomposed into a component explained by the pedigree-based kinship and a component based on the marker data. Consequently, the vector of phenotypic values is modeled as

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zt} + \mathbf{Zu} + \mathbf{e}$$

where vectors $\mathbf{t}$ and $\mathbf{u}$ comprise the random testcross values based on the pedigree and genomic kinship between DH lines, respectively. Both vectors are assumed to be independent and follow a normal distribution with $\mathbf{t} \sim N(0, \mathbf{K}\sigma_t^2)$ and $\mathbf{u} \sim N(0, \mathbf{U}\sigma_u^2)$, where $\mathbf{K}$ is the pedigree-based kinship matrix from Model 1 and $\mathbf{U}$ is the realized kinship matrix based on SNP marker data from Model 2. Vectors $\boldsymbol{\beta}$ and $\mathbf{e}$ and design matrices $\mathbf{X}$ and $\mathbf{Z}$ are defined as in Model 1.

Model SM

In plant breeding, the genomic kinship between DH lines is frequently estimated with an alternative measure of relatedness, the simple matching coefficient (SM, Sneath and Sokal 1973). For pairs of fully homozygous lines it is analogous to $1 - D$ where D is Rogers' distance (Rogers 1972), which has been shown to be linearly related to Malécot's coefficient of coancestry (Malécot 1948; Melchinger et al. 1991). Calculating the realized kinship matrix between the DH lines based on the simple matching coefficient gives Model SM, with the vector $\mathbf{y}$ being modeled as

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zs} + \mathbf{e}$$

where the vector $\mathbf{s}$ comprises testcross values assumed to be randomly distributed with $\mathbf{s} \sim N(0, \mathbf{S}\sigma_s^2)$, σ_s^2 is the testcross variance pertaining to Model SM and vectors $\boldsymbol{\beta}$ and $\mathbf{e}$ and design matrices $\mathbf{X}$ and $\mathbf{Z}$ are defined as in Model 1.

The $N \times N$ matrix $\mathbf{S}$ of realized kinship coefficients between DH lines was calculated from SNP data as follows. The matrix

$$\mathbf{S}_{\mathrm{SM}} = \frac{(\mathbf{W} - \mathbf{J}_{N\times M})(\mathbf{W} - \mathbf{J}_{N\times M})' + M\mathbf{J}_{N\times N}}{2M}$$

with $\mathbf{W}$ being the $N \times M$ design matrix assigning genotypes to phenotypes, $\mathbf{J}_{N\times M}$ and $\mathbf{J}_{N\times N}$ being a matrix of 1s with dimensions $N \times M$ and $N \times N$, respectively, and M being the number of markers, gives similarity coefficients analogous to the simple matching coefficient. Following Hayes and Goddard (2008) each element of the matrix $\mathbf{S}_{\mathrm{SM}}$ was transformed by subtracting the minimum value $s_{\min}$ of matrix $\mathbf{S}_{\mathrm{SM}}$ and standardizing by multiplying with $\frac{1}{1-s_{\min}}$. This leads to the matrix

$$\mathbf{S} = \frac{(\mathbf{W} - \mathbf{J}_{N\times M})(\mathbf{W} - \mathbf{J}_{N\times M})' + M\mathbf{J}_{N\times N} - 2Ms_{\min}\mathbf{J}_{N\times N}}{2M(1 - s_{\min})}$$

which was used as realized genomic kinship matrix in model SM.

In the Appendix, the dependency of the genetic variance components pertaining to Model 2 and Model SM is given. We also demonstrate the dependency of the genetic variance components of both models with the random regression model (Model RR) suggested by Meuwissen et al. (2001), where genomic testcross values are predicted based on individual SNP effects.

Variance components for all models were estimated with residual maximum likelihood (REML) using the ASReml software version 3.0 (Gilmour et al. 2009). Goodness of fit of the estimated models was assessed based on Akaike's information criterion $\mathrm{AIC} = -2L + 2k$ with L being the value of the log-likelihood function evaluated for the estimated parameter array and k the number of free parameters in the model (Akaike 1974). A likelihood ratio test according to Self and Liang (1987) was employed for comparison of Model 3 with the nested reduced Models 1 and 2.

Cross-validation

Fivefold cross-validation (CV) was used to assess the prediction accuracy of the different statistical models and sample sizes (Mosier 1951). The data set (DS) was divided into five mutually exclusive subsets, four of them formed the estimation set (ES) for fitting effects of the models and the fifth subset was used as test set (TS) for testcross value prediction. Different sampling strategies were employed to account for family structure in the DS. Within family sampling (CV-W) was performed for the entire DS ($N = 1{,}377$). In CV-W, testcross progenies of each family were subdivided into five subsets, four of them were assigned to the ES ($N_{ES} = 1{,}093$–$1{,}113$), one to the TS. With across family sampling (CV-A), the 36 families of the DS were divided into four subsets of seven and one subset of eight families. Thus, the ES comprised 28 or 29 families, and the TS comprised the remaining families not contained in the ES. Because the size of the families varied, the sample size of the ES also varied from $N_{ES} = 1{,}002$ to $N_{ES} = 1{,}172$. Assignment of genotypes to subsets was repeated ten times resulting in 50 different CV runs for each model.

To evaluate the effect of sample size on the accuracy of testcross prediction, the DS with 1,377 DH lines was divided into 2, 4 and 8 subsets resulting in an array of subsets of size $N = 688$, 344 and 172, respectively. The procedure was repeated 16 times for $N = 688$, 8 times for $N = 344$, and 4 times for $N = 172$ to create 32 subsets for each sample size. Each of them was analyzed with Model 1, 2 and 3. Fivefold CV was applied to the subsets. Sampling did not take into account the family structure of the testcross progenies.

In all CV scenarios, variance components needed for estimation of best linear unbiased predictors in the ES were derived from fitting the models to the DS. Average predictive abilities did not differ when variance components were re-estimated in the ES but the computational load was substantially greater. The population mean $\hat{\boldsymbol{\beta}}$ and the vectors $\hat{\mathbf{t}}$ and $\hat{\mathbf{u}}$ of length N were estimated from each ES. Testcross values of DH lines in the corresponding TS were predicted as $\hat{\mathbf{g}}_{TS} = \mathbf{X}_{TS}\hat{\boldsymbol{\beta}} + \mathbf{Z}_{TS}\hat{\mathbf{t}}$ (Model 1), $\hat{\mathbf{g}}_{TS} = \mathbf{X}_{TS}\hat{\boldsymbol{\beta}} + \mathbf{Z}_{TS}\hat{\mathbf{u}}$ (Model 2), and $\hat{\mathbf{g}}_{TS} = \mathbf{X}_{TS}\hat{\boldsymbol{\beta}} + \mathbf{Z}_{TS}\hat{\mathbf{t}} + Z_{TS}\hat{\mathbf{u}}$ (Model 3). Here, $\mathbf{X}_{TS}$ and $\mathbf{Z}_{TS}$ are design matrices where the number of rows equals the size of the TS and the number of columns equals one and N, respectively.

The correlation between observed and predicted testcross values ($\bar{r}(\mathbf{y}_{TS}, \hat{\mathbf{g}}_{TS})$) describes the predictive ability of a model and can be estimated directly from the data. The accuracy of a model, i.e. the correlation between true and predicted testcross values, can be approximated from the predictive ability as $r(\mathbf{g}_{TS}, \hat{\mathbf{g}}_{TS}) = r(\mathbf{y}_{TS}, \hat{\mathbf{g}}_{TS})/h$, where h is the square-root of the trait heritability (Dekkers 2007; Legarra et al. 2008). For both traits, broad sense heritability estimates were derived from variance component estimates of Model 1 as $\hat{h}^2 = \hat{\sigma}_t^2/(\hat{\sigma}_t^2 + \hat{\sigma}^2)$ according to standard quantitative genetic theory. From the ten replications of CV the mean accuracy was calculated. The standard error of the accuracy was calculated from the mean of the five mutually exclusive subsets of each replication as suggested by Luan et al. (2009). Significance of the difference in accuracy between models was tested with Student's paired t-test.

With each of the models, we computed the mean phenotypic testcross performance of the 10% best DH lines. For each of the ten CV runs, predicted testcross values from the five test sets were merged and the best 140 DH lines were selected based on their predicted testcross performance for grain yield. The observed testcross performance of the selected lines was averaged over the ten replications and differences of mean phenotypic testcross performance between models were tested with a Student's t-test.

To obtain predictive abilities for testcross performance within families, the four largest biparental families comprising 58–60 DH lines were analyzed individually. The expected kinship was close to 0.5 for all four families. Genomic testcross values were predicted with Model 2 and cross-validated with fivefold random sampling repeated ten times.

Results

Genotypic data

From the set of 732 polymorphic SNP markers 663 could be assigned to the ten maize chromosomes. For 69 markers the chromosomal position was unknown. The number of SNPs per linkage group ranged from 96 on chromosome 1 to 48 on chromosome 10. In the DH population, the average MAF was 0.19. 42% of the markers had a MAF $<$ 0.1. The number of polymorphic SNPs within the 36 families ranged from 78 to 600. The decline of LD in the DH population relative to the physical distance on the reference map (Schnable et al. 2009) is depicted in Fig. 1. As expected for an advanced cycle breeding population substantial long-range LD was detected. For the 31 parental inbred lines the magnitude of LD between pair-wise marker combinations is given in Supplemental Fig. 1.

Means and variance components

Adjusted testcross values of the 1,377 DH lines for GDY averaged across the seven locations ranged from 105.69 to

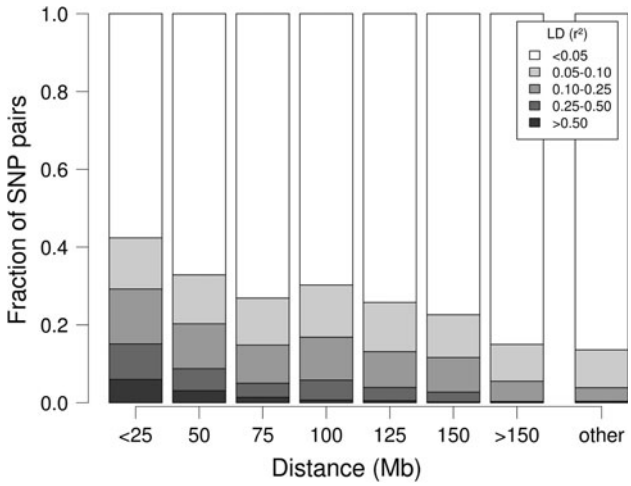

Fig. 1 Linkage disequilibrium between pair-wise marker combinations exhibiting significant LD ($p < 0.05$) within the same linkage group as a function of physical distance on the reference map for 1,377 DH progenies. Linkage disequilibrium of SNPs located on different linkage groups is displayed in the bar on the right-hand side

175.98 dt/ha and for GDC from 78.18 to 84.52%. Family means varied for GDY from 132.38 to 159.52 dt/ha and for GDC from 80.34 to 83.09%. Phenotypic correlations between the seven locations calculated from the testcross performance of the DH lines varied between 0.19 and 0.38 for GDY and between 0.32 and 0.64 for GDC. Models utilizing genomic information for calculation of the kinship matrix between DH lines showed a considerably smaller AIC than Model 1 where the kinship was exclusively modeled based on pedigree data (Table 1). Including a polygenic component in addition to marker information gave improved goodness of fit of Model 3 ($p < 0.001$) compared to Model 2 for both traits. In terms of goodness of fit, Model 2, Model SM, and Model RR did not differ as shown by their maximized log-likelihoods. The relationship between their genetic variance components are given in the Appendix. Dividing the estimated testcross variance component of Model 2 ($\hat{\sigma}_u^2$) by $8\sum_{m=1}^{M} p_m(1-p_m) = 8 \times 98.6 = 788.8$ gives variance component $\hat{\sigma}_m^2$ of Model RR. Multiplying the estimated testcross variance component of Model RR with $2M(1-s_{\min}) = 2 \times 732 \times 0.504 = 737.9$ gives variance components $\hat{\sigma}_s^2$ of Model SM.

Predictive abilities and accuracies of the different models

Predicted testcross values calculated from Models 2, SM, and RR were shifted by a constant but correlated with $r = 1.0$. Thus, the predictive abilities and accuracies derived from Model 2 are also representative for Models SM and RR and therefore we will report them jointly as results from Model 2. Predictive abilities obtained with CV-W, i.e. within family sampling, and CV-A, i.e. across family sampling, are shown for Models 1, 2, and 3 for both traits in Fig. 2. For GDY and CV-W, the average predictive ability ($\bar{r}(\mathbf{y}_{TS}, \hat{\mathbf{g}}_{TS})$) of Models 2 and 3 over the 50 CV runs was 0.66 and 0.68, respectively, which was higher ($p < 0.001$) than for Model 1 ($\bar{r}(\mathbf{y}_{TS}, \hat{\mathbf{g}}_{TS})$). Model 3 performed better than Model 2 ($p < 0.001$) but predicted testcross values were highly correlated ($r = 0.90$). Average predictive abilities of CV-A were reduced ($p < 0.001$) and much more variable compared to CV-W. For GDY, the average predictive ability for Model 1 was reduced to 0.11

Table 1 Estimates of variance components, their standard errors, mean (μ), log-likelihood (Log L) and Akaike's information criterion (AIC) for testcross progenies of 1,377 DH lines evaluated in seven locations estimated with five statistical models

Model	$\hat{\sigma}_t^2$	$\hat{\sigma}_u^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_m^2$	$\hat{\sigma}^2$	μ	Log L	AIC
Grain dry matter yield								
1	79.36 ± 19.49[a]				14.86 ± 10.3	148.14	−3,499.69	7,003
2		84.44 ± 10.50			34.92 ± 1.56	151.39	−3,343.46	6,691
3	36.92 ± 12.14	70.39 ± 9.46			14.37 ± 6.45	150.01	−3,318.45	6,643
SM			79.00 ± 9.86		34.92 ± 1.56	137.13	−3,343.46	6,691
RR[b]				0.107 ± 0.013	34.92 ± 1.56	148.46	−3,343.46	6,691
Grain dry matter content								
1	0.862 ± 0.22				0.1524 ± 0.12	81.51	−376.53	757
2		1.380 ± 0.15			0.2842 ± 0.01	81.43	−123.12	250
3	0.213 ± 0.09	1.282 ± 0.14			0.1673 ± 0.05	81.37	−113.39	233
SM			1.291 ± 0.14		0.2842 ± 0.01	81.43	−123.12	250
RR				0.0017 ± 0.0002	0.2842 ± 0.01	81.99	−123.12	250

[a] Standard error attached

[b] For details on Model RR see Appendix

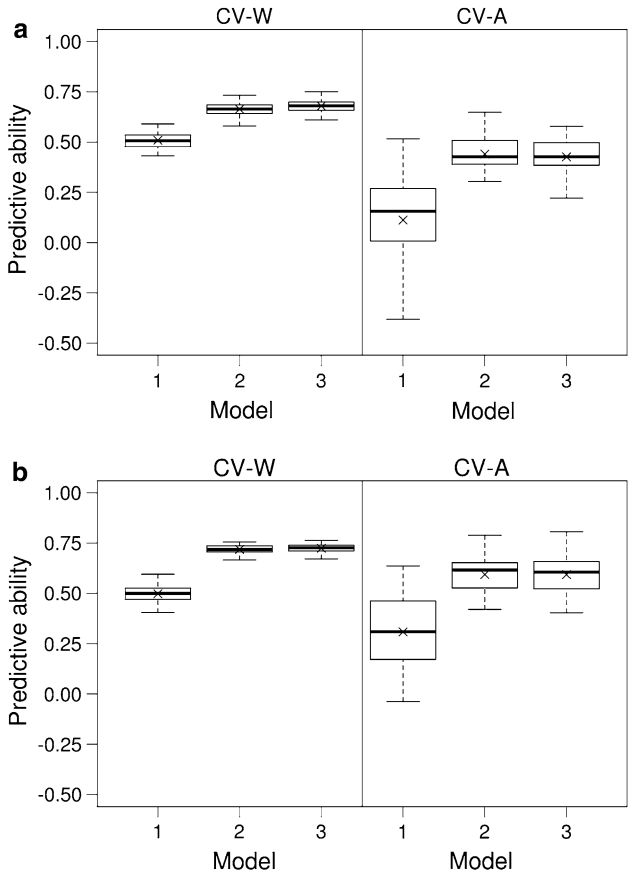

Fig. 2 Predictive abilities obtained with cross-validation with within family sampling (CV-W) and across family sampling (CV-A) for Models 1, 2 and 3 for traits grain dry matter yield (**a**) and grain dry matter content (**b**). The symbol × indicates the mean

and was lower than 0.43 for all models. Average predictive abilities for GDC were generally higher than for GDY in both CV schemes.

Average accuracies of CV-W and CV-A for the different models are shown for both traits in Table 2. Heritability estimates for the two traits were calculated from estimated variance components of Model 1 as $\hat{h}^2 = 0.84$ with a 95% confidence interval from 0.82 to 0.86 for GDY and $\hat{h}^2 = 0.85$ with a confidence interval from 0.83 to 0.86 for GDC.

The size of the data set had a significant effect ($p < 0.001$) on the accuracy of predicting testcross performance for grain yield with all models (Table 3). The reduction in accuracy was more pronounced for Models 2 and 3 than for Model 1, but performance of the models with genomic information was still significantly better.

Within each of the four biparental families the number of polymorphic markers varied from 116 to 212 (Table 4). Average predictive abilities differed significantly between families and varied from 0.24 to 0.54 for GDY and from 0.43 to 0.78 for GDC. Except for GDC in BP 3 predictive abilities were smaller than those obtained with the full data set.

Table 2 Average prediction accuracies and their standard errors derived from cross-validation with within family sampling (CV-W) and across family sampling (CV-A) for Models 1, 2 and 3 for traits grain dry matter yield (GDY) and grain dry matter content (GDC)

Model	CV-W		CV-A	
	GDY	GDC	GDY	GDC
1	0.56 ± 0.001[a]	0.54 ± 0.001	0.12 ± 0.039	0.33 ± 0.017
2	0.72 ± 0.002	0.78 ± 0.001	0.48 ± 0.012	0.64 ± 0.013
3	0.74 ± 0.002	0.79 ± 0.001	0.47 ± 0.017	0.64 ± 0.010

[a] Approximate standard errors attached

Table 3 Average prediction accuracies as well as their standard errors derived from cross-validation with random sampling for grain dry matter yield obtained with decreasing sample sizes ($N = 1{,}377$, 688, 344, 172) for Models 1, 2 and 3

Model	Sample size (N)			
	1,377	688	344	172
1	0.55 ± 0.001[a]	0.53 ± 0.002	0.51 ± 0.004	0.43 ± 0.005
2	0.72 ± 0.002	0.67 ± 0.002	0.62 ± 0.002	0.51 ± 0.004
3	0.74 ± 0.002	0.69 ± 0.002	0.63 ± 0.002	0.53 ± 0.005

[a] Approximate standard errors attached

Table 4 Average predictive abilities obtained with Model 2 and cross-validation with random sampling for the four largest biparental families for grain dry matter yield (GDY) and grain dry matter content (GDC)

Family	Sample size	M[a]	Predictive ability	
			GDY	GDC
BP 1	58	116	0.24 ± 0.037	0.43 ± 0.030
BP 2	60	132	0.45 ± 0.034	0.55 ± 0.026
BP 3	60	162	0.53 ± 0.036	0.78 ± 0.015
BP 4	59	212	0.54 ± 0.030	0.52 ± 0.024

[a] Number of polymorphic markers

Discussion

The current knowledge about the accuracy of predicting phenotypic performance from genome-wide random SNP markers has mainly been inferred from computer simulations (Bernardo and Yu 2007; Zhong et al. 2009). Recently published results from biparental maize and Arabidopsis populations (Lorenzana and Bernardo 2009) and a panel of diverse wheat and maize inbred lines (Crossa et al. 2010), have given first indications that breeding for complex traits

might benefit from incorporating genomic information in the selection process. However, these studies do not address prediction of testcross performance at the level of advanced breeding populations. In this study, we investigated the accuracy of predicting testcross performance of DH lines of maize derived from 36 crosses for two important quantitative traits, grain dry matter yield and content. Based on results from quantitave trait loci (QTL) mapping studies in large biparental populations (Melchinger et al. 1998; Schön et al. 2004) we inferred that the genetic architecture of these two traits can be approximated with Fisher's infinitesimal model. For traits regulated by a large number of genes with small effects and populations with strong long-range LD, mixed effects models have been shown to perform equally well with respect to prediction accuracies as Bayesian methods (Lorenzana and Bernardo 2009; Zhong et al. 2009; Piepho 2009; Crossa et al. 2010). These findings could be corroborated for this data set. Variable selection models, e.g. BayesB (Meuwissen et al. 2001), did not outperform Model 2 (Wimmer et al., unpublished).

When comparing the performance of the respective prediction models across traits, inferences must be based on prediction accuracies calculated from the ratio of the predictive abilities obtained from CV and the square-root of the trait heritability (Dekkers 2007). As pointed out by Piepho and Möhring (2007), estimation of the trait heritability in an advanced cycle breeding population is not straight forward if genotypes are not independent and the kinship matrix is used for modeling genetic effects. Further research is needed to address the problem of calculating progeny-mean heritabilities in kinship-based models. We estimated trait heritability using the standard formula $h^2 = \sigma_t^2/\left(\sigma_t^2 + \sigma^2\right)$ and variance component estimates derived from pedigree-based Model 1 being aware that these might lead to inflated estimates of h^2 (Piepho and Möhring 2007). We chose this estimate because it resulted in conservative estimates for the prediction accuracies.

Accuracies obtained with pedigree and genome-wide marker data

Independent of the CV scheme employed, the models based on genome-wide marker data performed significantly better than the model based on pedigree data alone. As pointed out by Goddard (2008), for predicting the magnitude of this increase in accuracy, the effective number of segregating loci in the population under study is most relevant. In an advanced cycle breeding population with small effective population size and doubled haploid lines mostly generated from biparental crosses, extensive long-range LD is expected and was shown for the experimental material under study (Fig. 1). Consequently, the variation in realized genetic relationship among DH lines sharing the same pedigree is high (Fig. 3), leading to an increase in prediction accuracy for models using genomic data. For 264 DH lines of a randomly chosen test set from CV-W, Fig. 4 illustrates predicted testcross values of DH lines for grain yield derived from Models 1 and 2 relative to their respective family mean calculated from adjusted means of the full data set. While in Model 1 testcross performance of

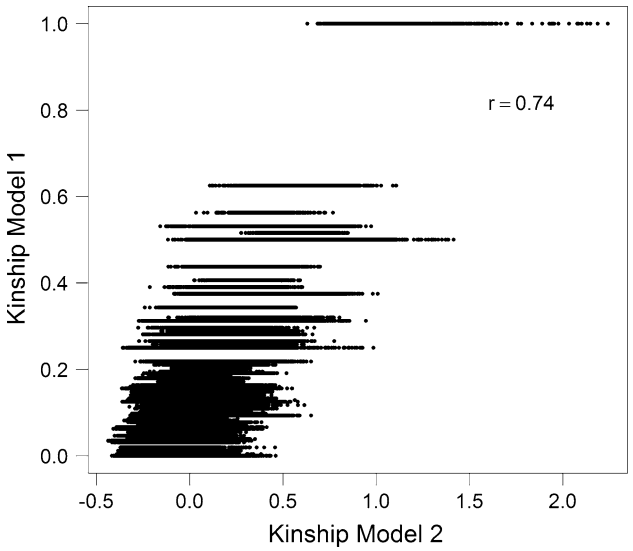

Fig. 3 Estimates of expected kinship coefficients between DH lines based on pedigree data (Matrix **K**, Model 1) against realized kinship based on marker data (Matrix **U**, Model 2). The correlation coefficient (r) between the values of both kinship matrices is displayed in the graph

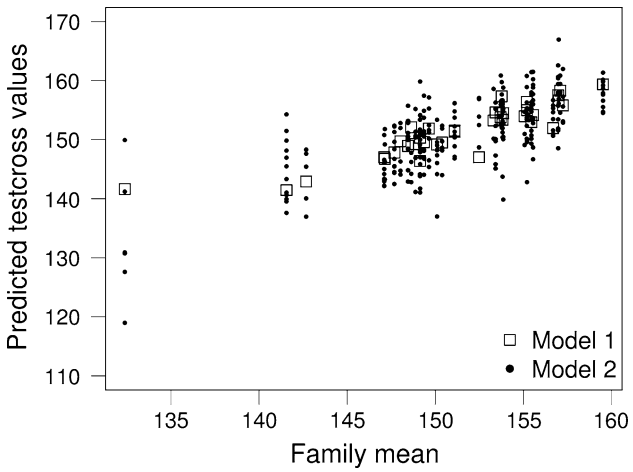

Fig. 4 Predicted testcross values obtained from one random test set of within family cross-validation (CV-W). Testcross values predicted with Model 1 and Model 2 are plotted against their respective family means calculated from adjusted means of the full data set for grain dry matter yield

all DH lines derived from the same cross is predicted by the same value, variation of testcross values predicted with Model 2 is large within each of the 36 families leading to a higher prediction accuracy.

Including genome-wide marker and pedigree information in the model (Model 3) improved prediction of testcross performance significantly but only to a small extent and only with CV-W. A modest improvement of predictive abilities of models including pedigree information in addition to marker data was also observed by Crossa et al. (2010). Goddard (2008) pointed out that including a polygenic term in the model might be beneficial for capturing the effects of alleles with low frequency. However, the relative performance of Model 3 was equal compared to Model 2 when predictions were calculated for distantly related DH lines in CV-A (Table 2). If high-quality pedigree and genome-wide marker data are available for all individuals in model training and prediction, Model 3 is not expected to consistently outperform Model 2 with mixed effects models due to the redundancy of the data.

Our conclusions on the relative performance of the three models were also confirmed when comparing the mean phenotypic testcross performance of the 10% best lines selected based on their predicted testcross performance for grain yield. Lines selected based on predictions from Models 2 and 3 performed significantly better than those selected based on Model 1 with both CV schemes, but Model 3 did not have an advantage over Model 2. However, further research is needed to assess the relative performance of Models 2 and 3 when pedigree and marker data are not redundant but complementary or show different levels of precision.

Comparison of cross-validation schemes

The prediction accuracy differed greatly depending on the stratification procedure employed in CV. CV-W yielded high average accuracies for both traits, when the genetic relationship between DH lines was modeled based on genomic data ($\bar{r}(\mathbf{g}_{TS}, \hat{\mathbf{g}}_{TS}) \geq 0.72$). These high values are in accordance with analytical and computer simulation results presented by Hayes et al. (2009b) who also showed high prediction accuracies within families as compared to random mating populations. The high accuracies obtained with CV-W are the result of high levels of relatedness between DH lines in the estimation and the test sets and long-range haplotype blocks within families leading to high LD between markers and QTL affecting the trait under study. Therefore, the given accuracies must be considered valid only for predictions of the performance of close relatives of the material under study.

In contrast, average prediction accuracies were significantly lower with CV-A. As compared to CV-W, the relative contribution of relatedness to the prediction accuracy is expected to decrease in CV-A and the relative importance of LD between markers and QTL is expected to increase with models utilizing genomic information. The low average accuracies obtained in CV-A with Model 1 ($\bar{r}(\mathbf{g}_{TS}, \hat{\mathbf{g}}_{TS}) \leq 0.33$) indicated that in the population under study, families derived from different crosses were only distantly related by pedigree. Prediction accuracies for Models 2 and 3 also decreased with CV-A as compared to CV-W, but not as severely as with Model 1. Thus, the reduced level of relatedness had a strong effect on prediction even for models using genomic information, but prediction accuracies remained at a medium level, indicating that substantial LD between the markers and QTL was captured by the markers. It needs to be noted, however, that accuracies varied considerably more between the 50 CV-A test sets as compared to CV-W, presumably as the result of the highly variable degree of relatedness between the DH lines in the estimation and the corresponding test sets, indicating that relatedness was still a component of the prediction accuracy. On the other hand, the larger variation of accuracies in CV-A might also result at least partly from the more variable sample sizes of the ES and TS in CV-A as compared to CV-W. Accuracies for dry matter content in CV-A were significantly higher than for grain yield, which was surprising because heritability estimates were quite similar for the two traits. This could be an indication that the trait GDC is controlled by fewer genes with larger effects which are better captured by the markers as compared to GDY.

As suggested by Luan et al. (2009), we calculated the regression of the observed on the predicted testcross performance in the 50 test sets of CV-W and CV-A. If the regression deviates from 1 this can be interpreted as a bias in predicted testcross values. A regression coefficient <1 implies inflation, a coefficient >1 deflation of predicted testcross values relative to their observed phenotypes. In CV-W, mean regression coefficients ranged from 0.94 (Model 2) to 1.02 (Model 1) indicating only a modest inflation of predicted testcross values for Model 2. In CV-A, however, regression coefficients ranged from 0.20 (Model 1) to 1.07 (Model 2) reflecting a strong upward bias in predicted testcross values for Model 1.

Accuracies in biparental families

Analytical results and computer simulations showed that prediction accuracies that can be obtained with full sib families are substantially higher than those for random mating populations because allele effects are estimated more accurately and the effective number of independently segregating loci controlling the phenotype is reduced (Hayes et al. 2009b). In addition, when model training is

performed across families with relatively low marker densities, inconsistent linkage phases between SNP markers and QTL might reduce the prediction accuracy in a population-wide approach. Therefore, we assessed genomic predictive abilities obtained with Model 2 and random sampling CV individually for four biparental families. The predictive abilities were higher for those families with more polymorphic markers indicating that marker coverage might have been a limiting factor in prediction of testcross values. Prediction accuracies were tremendously variable between the 50 CV runs within each family, most likely a result of the small sample size. Considering the substantially higher and much less variable predictive abilities obtained when using the full data set and CV-W, we follow Jannink et al. (2010) in their argumentation that it does not seem appropriate to perform model training within individual families not taking into account information from related families or population-wide LD.

Modeling the kinship of DH lines

The estimation of genomic relatedness was one of the first applications of molecular markers in plant breeding and has been successfully used for management of heterotic pools and diversity analyses. Many different measures of relatedness have been proposed for quantifying the kinship between pairs of individuals (Reif et al. 2005). In maize breeding, estimates of kinship between fully homozygous inbred lines are frequently calculated from the proportion of shared marker alleles corrected for the average proportion of alleles alike in state between unrelated individuals in the population under study (Bernardo 1993). In Model SM we followed a similar approach, but instead of correcting with the proportion of alleles alike in state estimated from unrelated individuals, we used the minimum value of alleles shared between DH lines ($s_{\min}$) as suggested by Hayes and Goddard (2008). The reason was that subtracting the average proportion of alleles alike in state determined from unrelated individuals can lead to the violation of the assumption of the variance–covariance matrix $\mathbf{S}$ being positive semidefinite. As has been shown in the Appendix, the choice of the correction factor in Model SM will affect the estimated variance components in a predictable form and predicted testcross values from Model SM will be shifted in scale but ranked identically when compared to Model 2 and Model RR.

In Model 2, the kinship between DH lines is estimated from genome-wide covariances of allele counts which can be interpreted as deviations of allele sharing from that expected for unrelated individuals. As pointed out by Habier et al. (2007), correction of allele counts with their expected values subtracts the same constant for all individuals in the population under study and thus only leads to a scale-shift. Analogously as for Model SM, testcross values predicted with Model 2 are shifted in scale but ranked identically when compared to Model RR. Thus, for estimation of the realized genetic relationship matrix based on genome-wide marker data, we decided to focus on Model 2, because if estimators of kinship coefficients are linear combinations of one another their prediction accuracies are identical. One advantage of Model 2 is, that with the high marker densities expected in future analyses the computational burden of Model 2 is substantially reduced compared to Model RR.

Application of genomic selection in maize breeding

The results from this study give first insights into the potential of genome-based prediction of testcross performance in maize. The population under study was chosen to be representative of a typical advanced cycle breeding population with respect to family structure, allelic diversity, and extent of LD. The marker density ($M = 732$) warranted good genome coverage across the entire population and within families. Family means of the 36 crosses varied significantly for grain yield (132.4–159.5 dt/ha) and grain dry matter content (80.3–83.1%). The size of the training population was large ($N = 1{,}377$). However, prediction accuracies obtained with the given experimental material can only be taken as points of reference and must be interpreted within their specific context. The number of test environments was rather high for a first testcross evaluation and the trial design was completely balanced. Accuracies obtained with CV-W of Model 2 are indicative for predicting the testcross performance of DH lines without phenotypes from the same cycle of selection and with close relatives in the training population. Assuming that a certain degree of relatedness generally persists in breeding populations, accuracies obtained with CV-A of Model 2 should provide a realistic estimate for the prediction accuracy of DH lines not necessarily closely related to lines in the training population, but from the same cycle of selection. Prediction of performance in subsequent cycles of selection is yet to be investigated. The effects of recombination and selection, marker × year interactions, or the use of a different tester will decrease prediction accuracies. Thus, constant updating of the prediction model will be essential for genomic selection to become useful in the breeding process. The design of the training population with respect to sample size, number and type of crosses, number of progenies per cross, number of test environments and testers still need to be investigated. In this study, estimation sets of size $N_{ES} = 550$ (data set $N = 688$) still performed quite well with prediction accuracies larger than 0.67 for grain yield, most likely the result of the high trait heritability. It needs to be kept in mind, however,

that with decreasing sample size prediction accuracies became more variable, which might be even more pronounced when predicting the performance of distantly related individuals.

To our knowledge this is the first experimental study in plants reporting prediction accuracies for genomic testcross values estimated at the population level. For the given marker density and sample size, accuracies were high especially when compared to values reported in the literature for random mating populations in animal breeding (Hayes et al. 2009a; Moser et al. 2010). Thus, our results were encouraging with respect to genome-based prediction of the genetic value of untested lines in advanced cycle breeding populations and the implementation of genomic selection in the breeding process. However, many questions remain concerning the application of genomic selection in plant breeding. One of the crucial points will be to quantify the relative contributions of relatedness and LD between markers and functional polymorphisms to the prediction accuracy in highly structured plant populations.

Acknowledgments This research was funded by the German Federal Ministry of Education and Research (BMBF) within the AgroClustEr "Synbreed—Synergistic plant and animal breeding" (FKZ: 0315528A).

Appendix

In this Appendix we show the functional dependency of the genomic variance components pertaining to Model 2, Model SM and the random regression model suggested by Meuwissen et al. (2001) denoted Model RR.

All three models can be seen as special cases of a more general model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{W} - \mathbf{Q})\mathbf{m} + \mathbf{e}$$

where $\boldsymbol{\beta}$ is a vector of fixed effects and $\mathbf{X}$ is a design matrix assigning fixed effects to the phenotypes, $\mathbf{W}$ is an $N \times M$ design matrix assigning M SNP marker genotypes coded 0 or 2 to N phenotypes. Random marker effects in vector $\mathbf{m}$ are assumed to follow a normal distribution with $\mathbf{m} \sim N(0, \mathbf{I}\sigma_m^2)$, where $\mathbf{I}$ is an identity matrix and σ_m^2 denotes the proportion of the testcross variance contributed by each individual SNP marker. $\mathbf{Q}$ is an $N \times M$ matrix composed of M uniform column vectors $\mathbf{Q} = \{q_1\mathbf{c}, q_2\mathbf{c}, \cdots, q_M\mathbf{c}\}$ where q_m is a scalar correction term for marker m and $\mathbf{c}$ is a column vector of length N containing only 1s.

The estimated SNP effects and the corresponding variance component σ_m^2 in the general model are unaffected by the choice of $\mathbf{Q}$, since subtracting the correction terms shifts the intercept, but does not change the slope of the regression of the phenotype on each SNP (Habier et al. 2007; Piepho 2009). For Model RR, $\mathbf{Q}$ is the null matrix, but the same variance components as estimated from Model RR will be received with any other $\mathbf{Q}$.

For Model 2 the kinship between lines i and j is modeled by the matrix $\mathbf{U}$ calculated as

$$\mathbf{U} = \frac{(\mathbf{W} - \mathbf{P})(\mathbf{W} - \mathbf{P})'}{8\sum_{m=1}^{M} p_m(1 - p_m)}$$

and the variance–covariance matrix of the phenotype vector $\mathbf{y}$ can be written as

$$\mathbf{V}_{\text{Model2}} = \frac{(\mathbf{W} - \mathbf{P})(\mathbf{W} - \mathbf{P})'}{8\sum_{m=1}^{M} p_m(1 - p_m)}\sigma_u^2 + \mathbf{I}\sigma^2$$

Using $\mathbf{Q} = \mathbf{P}$ in the generalized model, the variance–covariance matrix of the phenotype vector $\mathbf{y}$ is

$$\mathbf{V}_{\text{P}} = (\mathbf{W} - \mathbf{P})(\mathbf{W} - \mathbf{P})'\sigma_m^2 + \mathbf{I}\sigma^2$$

Hence, $\mathbf{V}_{\text{Model2}} = \mathbf{V}_P$ if

$$\sigma_u^2 = \sigma_m^2 \times 8\sum_{m=1}^{M} p_m(1 - p_m)$$

In Model SM, the matrix $\mathbf{S}$ can be written as

$$\mathbf{S} = \frac{(\mathbf{W} - \mathbf{J}_{N\times M})(\mathbf{W} - \mathbf{J}_{N\times M})'}{2M(1 - s_{\min})} + \frac{M - 2Ms_{\min}}{2M - 2Ms_{\min}}\mathbf{J}_{N\times N}$$

The second term reflects a constant, which is added to all elements of the matrix. This is equivalent to a constant random block effect and thus fully confounded with the fixed intercept (Piepho et al. 2008; Williams et al. 2006). Hence, ignoring the second term will not affect the estimated variance components.

The numerator of the first term of $\mathbf{S}$ is a special case of the numerator of the generalized model with $\mathbf{Q} = \mathbf{J}_{N\times M}$, i.e. assuming $q_m = 1$ for all loci, which is equivalent to the assumption of allele frequency $p_m = 0.5$ for all SNPs in Model 2. Using the same argument as above,

$$\mathbf{V}_{\text{SM}} = \frac{(\mathbf{W} - \mathbf{J}_{N\times M})(\mathbf{W} - \mathbf{J}_{N\times M})'}{2M(1 - s_{\min})}\sigma_s^2 + \mathbf{I}\sigma^2$$

For the special case $\mathbf{Q} = \mathbf{J}_{N\times M}$ in the general model, the variance–covariance matrix of the phenotype vector $\mathbf{y}$ is

$$\mathbf{V}_J = (\mathbf{W} - \mathbf{J}_{N\times M})(\mathbf{W} - \mathbf{J}_{N\times M})'\sigma_m^2 + \mathbf{I}\sigma^2$$

and $\mathbf{V}_{\text{SM}} = \mathbf{V}_J$ if

$$\sigma_s^2 = \sigma_m^2 \times 2M(1 - s_{\min}).$$

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Auto Control 19:716–723

Bernardo R (1993) Estimation of coefficient of coancestry using molecular markers in maize. Theor Appl Genet 85:1055–1062
Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN
Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. Crop Sci 47:1082–1090
Crossa J, de los Campos G, Pérez P, Gianola D, Burgueño J, Araus JL, Makumbi D, Singh RP, Dreisigacker S, Yan J, Arief V, Banzinger M, Braun HJ (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186:713–724
Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. J Anim Breed Genet 124:331–341
Foulkes AS (2009) Applied statistical genetics with R: for population-based association studies, 1st edn. Springer, New York
Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) ASReml user guide release 3.0. VSN International Ltd. Hemel Hempstead
Goddard ME (2008) Genomic selection: prediction of accuracy and maximization of long term response. Genetica 136:245–257
Grubbs FE (1950) Sample criteria for testing outlying observations. Ann Math Stat 21:27–58
Habier D, Fernando RL, Dekkers JCM (2007) The impact of genetic relationship information on genome-assisted breeding values. Genetics 177:2389–2397
Habier D, Tetens J, Seefried FR, Lichtner P, Thaller G (2010) The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genet Sel Evol 42:5
Hayes BJ, Goddard ME (2008) Technical note: prediction of breeding values using marker-derived relationship matrices. J Anim Sci 86:2089–2092
Hayes BJ, Daetwyler HD, Bowman P, Moser G, Tier B, Crump R, Khatkar M, Raadsma HW, Goddard ME (2009a) Accuracy of genomic selection: comparing theory and results. Proc Assoc Adv Anim Breed Genet 18:34–37
Hayes BJ, Visscher PM, Goddard ME (2009b) Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res 91:47–60
Henderson CR (1984) Applications of linear models in animal breeding. University of Guelph, Guelph
Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231
Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 9:166–177
Legarra A, Robert-Granie C, Manfredi E, Elsen J (2008) Performance of genomic selection in mice. Genetics 180:611–618
Lorenzana R, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120:151–161
Luan T, Wooliams JA, Lien S, Kent M, Svendsen M, Meuwissen THE (2009) The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. Genetics 183:1119–1126
Malécot G (1948) Les mathématiques de l'hérédité. Masson et Cie, Paris
Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci 31:669–678
Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403
Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829
Moser G, Khatkar MS, Hayes BJ, Raadsma HW (2010) Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP markers. Genet Sel Evol 42:37
Mosier CI (1951) I. Problems and designs of cross-validation 1. Educ Psychol Measurement 11:5–11
Piepho HP (2009) Ridge regression and extensions for genomewide selection in maize. Crop Sci 49:1165–1176
Piepho HP, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888
Piepho HP, Richter C, Williams E (2008) Nearest neighbor adjustment and linear variance models in plant breeding trials. Biom J 50:164–189
Reif JC, Melchinger AE, Frisch M (2005) Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. Crop Sci 45:1–7
Röber F, Gordillo G, Geiger HH (2005) In vivo haploid induction in maize—performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50:275–283
Rogers JS (1972) Measures of genetic similarity and genetic distance. In: Studies in genetics VII. University of Texas, Austin, pp 145–153
Schaeffer LR (2006) Strategy for applying genome-wide selection in dairy cattle. J Anim Breed Genet 123:218–223
Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L et al (2009) The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112–1115
Schnell FW (1965) Die Covarianz zwischen Verwandten in einer genorthogonalen Population. I. Allgemeine Theorie. Biom Z 7:1–49
Schön CC, Utz HF, Groh S, Truberg B, Openshaw S, Melchinger AE (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485–498
Self SG, Liang KY (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Ass 82:605–610
Sneath PH, Sokal RR (1973) Numerical taxonomy: the principles and practice of numerical classification. Freeman, San Francisco, CA
Stuber CW, Cockerham CC (1966) Gene effects and variances in hybrid populations. Genetics 54:1279–1286
VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423
VanRaden PM, Tassell CV, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92:16–24
Williams ER, John JA, Whitaker D (2006) Construction of resolvable spatial row–column designs. Biometrics 62:03–108
Zhong S, Dekkers JCM, Fernando RL, Jannink JL (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: A barley case study. Genetics 182:355–364