Theor Appl Genet (2010) 120:451–461
DOI 10.1007/s00122-009-1208-x

ORIGINAL PAPER

Prediction of hybrid performance in maize using molecular markers and joint analyses of hybrids and parental inbreds

Tobias A. Schrag · Jens Möhring · Albrecht E. Melchinger · Barbara Kusterer · Baldev S. Dhillon · Hans-Peter Piepho · Matthias Frisch

Received: 31 May 2009 / Accepted: 21 October 2009 / Published online: 15 November 2009

Abstract The identification of superior hybrids is important for the success of a hybrid breeding program. However, field evaluation of all possible crosses among inbred lines requires extremely large resources. Therefore, efforts have been made to predict hybrid performance (HP) by using field data of related genotypes and molecular markers. In the present study, the main objective was to assess the usefulness of pedigree information in combination with the covariance between general combining ability (GCA) and per se performance of parental lines for HP prediction. In addition, we compared the prediction efficiency of AFLP and SSR marker data, estimated marker effects separately for reciprocal allelic configurations (among heterotic groups) of heterozygous marker loci in hybrids, and imputed missing AFLP marker data for marker-based HP prediction. Unbalanced field data of 400 maize dent × flint hybrids from 9 factorials and of 79 inbred parents were subjected to joint analyses with mixed linear models. The inbreds were genotyped with 910 AFLP and 256 SSR markers. Efficiency of prediction (R^2) was estimated by cross-validation for hybrids having no or one parent evaluated in testcrosses. Best linear unbiased prediction of GCA and specific combining ability resulted in the highest efficiencies for HP prediction for both traits ($R^2 = 0.6$–0.9), if pedigree and line per se data were used. However, without such data, HP for grain yield was more efficiently predicted using molecular markers. The additional modifications of the marker-based approaches had no clear effect. Our study showed the high potential of joint analyses of hybrids and parental inbred lines for the prediction of performance of untested hybrids.

Communicated by A. Charcosset.

Contribution to the special issue 'Heterosis in Plants'.

T. A. Schrag · A. E. Melchinger (✉) · B. S. Dhillon
Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany
e-mail: melchinger@uni-hohenheim.de

J. Möhring · H.-P. Piepho
Bioinformatics Unit, Institute for Crop Production and Grassland Research, University of Hohenheim, 70599 Stuttgart, Germany

B. Kusterer
HYBRO Saatzucht GmbH & Co. KG, Kleptow Nr. 53, 17291 Schenkenberg, Germany

M. Frisch
Institute of Agronomy and Plant Breeding II, Justus-Liebig-University Giessen, 35392 Giessen, Germany

Introduction

For hybrid variety development, maize breeders continuously develop a large number of inbred lines. This has been facilitated and accelerated in recent years by establishing the doubled haploid technology (Schmidt 2004; Seitz 2005). With an increase in the number of inbreds, the number of crosses between lines from different heterotic pools grows very rapidly and their field evaluation requires large resources. Thus, in practice only a small proportion of all possible experimental hybrids are evaluated in field trials. Identification of promising inter-pool hybrids without having them tested in the field has been attempted by prediction of hybrid performance (HP) utilizing field trial data available from related crosses. Promising results were obtained with best linear unbiased prediction (BLUP) of HP, using mixed linear models for the analysis of phenotypic trait data together with coancestry coefficients estimated from pedigree records or marker data (Bernardo 1994).

Mixed linear models for the joint analysis of field data from hybrids and their parental inbreds allow adjustment for trial and environmental effects and facilitate the estimation of mid-parent heterosis across several field experiments. In such a joint analysis, covariance between the per se performance and the general combining ability (GCA) of the parental lines can be considered in the model. In addition, covariance among parental lines can be modeled with pedigree-based coancestry coefficient matrices (Piepho et al. 2008). The gain of information from related genotypes is expected to improve the prediction, especially of hybrids between parents for which no testcross data are available (Schrag et al. 2009). However, the degree of improvement in HP prediction by considering the covariance between GCA and line per se performance in combination with pedigree-based relationship matrices has not been investigated so far.

Since the advent of molecular markers, various marker-based measures have been used to predict HP in maize. These measures are genetic distance (Lee et al. 1989; Bernardo 1992; Charcosset and Essioux 1994), hybrid value (Dudley et al. 1991) and predicted specific combining ability (SCA) (Charcosset et al. 1998). Another approach to make use of molecular markers for HP prediction is support vector machine regression (Maenhout et al. 2010). In BLUP analyses of field data, Bernardo (1999) included marker data to account for quantitative trait loci (QTL). However, only marginal improvements were observed if compared with a model using field data alone. Linear regression on marker-based estimates of genotypic values was used by Vuylsteke et al. (2000) to predict HP and SCA of inter-pool crosses in maize. This approach was extended and validated for the prediction of grain yield (GY) and grain dry matter content (GDMC) in dent × flint factorial crosses using bi-allelic amplified fragment length polymorphism (AFLP) markers by Schrag et al. (2006, 2007).

Multi-allelic marker systems, such as simple sequence repeat (SSR) markers, play an important role in plant breeding (Karakousis et al. 2003). They are expected to carry higher information content per marker locus than bi-allelic markers and thereby enhance the probability of distinguishing different functional alleles for estimation of marker effects. However, the number of observations per allelic configuration of marker loci in hybrids is reduced. This leads to the question, whether efficiency of HP prediction based on a SSR marker data set is higher than with an AFLP marker data set, assuming that both are dimensioned with comparable expenses.

Previous studies ignored whether alleles that were identical in state were contributed by the dent or by the flint parent (Vuylsteke et al. 2000; Schrag et al. 2007). For a given marker locus of a hybrid, the reciprocal allelic configurations of heterozygous marker loci were therefore not distinguished for estimation of marker effects. This model assumes that a marker allele is coupled with the same QTL allele in both heterotic pools. However, the European flints were introduced into Europe more than 500 years ago (Rebourg et al. 2003), whereas the dent lines in Central Europe were derived from US dent lines during the past 50 years (Stich et al. 2005). Hence, dent and flint lines have been genetically separated for a long period and, therefore, the linkage disequilibria between markers and QTL may differ between both heterotic pools (Charcosset and Essioux 1994). This raises the question, whether efficiency of HP prediction increases by estimating the marker effects separately for reciprocal allelic configurations of heterozygous marker loci in hybrids.

Markers with missing observations cannot be used for prediction approaches based on the total contribution of selected markers (TCSM, Vuylsteke et al. 2000) or multiple linear regression (MLR, Schrag et al. 2007). As the number of parental genotypes increases, this problem becomes very severe due to the increasing probability of at least one dropout for a given marker. One possibility is to develop prediction models that accept missing observations, such as the TEAM approach (Schrag et al. 2007). A more general approach is to estimate missing marker data from the observed genotypes of tightly linked markers (Balding 2006; Roberts et al. 2007). This promises to be a simple and effective solution for prediction approaches that otherwise cannot handle missing marker data. However, the usefulness of such data imputation for marker-based prediction of HP has not been studied so far.

The goal of our study was to investigate marker-based prediction of the performance of inter-pool maize hybrids using joint analyses of hybrids and parental lines. The main objective was to study the advantage of including pedigree-based estimates of covariance among relatives as well as covariance between GCA and per se performance of parental lines in the mixed linear model. In addition, we compared the prediction efficiency using SSR instead of AFLP marker data, investigated the efficiency of prediction using marker effects that were estimated separately for reciprocal allelic configurations of heterozygous marker loci in hybrids, and analyzed the benefit of imputing missing AFLP marker data for marker-based HP prediction.

Materials and methods

Phenotypic and pedigree data

Field data and pedigree records were used as described by Schrag et al. (2009). Briefly, nine factorial mating designs

of dent × flint maize hybrids were evaluated each in a 1-year multi-location experiment. This resulted in 54 experiment by location combinations, designated as single trials, and involved 6 years and 11 locations in total. Between 5 and 9 check hybrid varieties were included in each experiment. By combining these data, an unbalanced factorial between 47 dent and 32 flint inbred parents was generated, comprising 400 tested hybrids. The 79 inbred parents and 41 additional lines were evaluated for their per se performance in 11 single trials, involving 3 years and 5 locations. All trials were carried out using adjacent α-designs with two-row plots, five plots per block, and two to three replicates. GY was recorded in Mg ha^{-1} adjusted to 155 g kg^{-1} grain moisture, while GDMC was recorded in percent.

Mixed linear model analysis

Data for GY and GDMC from all experiments were analyzed with mixed linear models. The 41 additional lines in the line per se experiment, which were not used as parents in factorials matings, were treated as checks. These and the check varieties in the experiments of the hybrids were used to adjust for single trial effects, but were not used to estimate genetic variances. Dummy variables were used to divide genotypes into hybrids, dent and flint parental lines, and checks. The model can be described in the syntax of Patterson (1997) by

$$\begin{aligned} & Y + L + Y \cdot L + C : T \cdot Y \cdot L + R \cdot T \cdot Y \cdot L \\ & \quad + B \cdot R \cdot T \cdot Y \cdot L + C \cdot L + C \cdot Y + C \cdot Y \cdot L \\ & \quad + GCA_{dent} + GCA_{flint} + SCA + line_{dent} + line_{flint} \\ & \quad + GCA_{dent} \cdot L + GCA_{flint} \cdot L + SCA \cdot L + line_{dent} \cdot L \\ & \qquad + line_{flint} \cdot L \\ & \quad + GCA_{dent} \cdot Y + GCA_{flint} \cdot Y + SCA \cdot Y + line_{dent} \cdot Y \\ & \qquad + line_{flint} \cdot Y \\ & \quad + GCA_{dent} \cdot Y \cdot L + GCA_{flint} \cdot Y \cdot L + SCA \cdot Y \cdot L \\ & \qquad + line_{dent} \cdot Y \cdot L + line_{flint} \cdot Y \cdot L \end{aligned} \tag{1}$$

where fixed and random effects were separated by a colon and interactions between two effects were denoted by a dot between the main effects. Main effects for the factors year (Y), location (L), and check (C) were treated as fixed effects, which allowed accounting for performance differences between single trials. Genotypic effects, all interactions, and effects of trials (T), replicates (R) within trials, and incomplete blocks (B) within replicates were treated as random. Genotypic effects of the factorial crosses were partitioned into GCA effects of the parental dent and flint inbreds, and SCA effects of the crosses (Gardner and Eberhart 1966). It was assumed that the vector of GCA effects was normally distributed with variances $\boldsymbol{A}_1\sigma^2_{GCAdent}$ and $\boldsymbol{A}_2\sigma^2_{GCAflint}$ for the dent and flint pools, respectively, where $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ were the additive relationship matrices for genotypes of the respective pools. They were computed from coefficients of coancestry among the inbred lines (Bernardo 2002). The vector of SCA effects was assumed to have variance $\boldsymbol{D}\sigma^2_{SCA}$, where $\boldsymbol{D}$ was a matrix with elements equal to the product of parental coefficients of coancestry. Inbreds were divided into dent ($line_{Dent}$) and flint ($line_{Flint}$) lines with variances $\boldsymbol{A}_1\sigma^2_{PERSEdent}$ and $\boldsymbol{A}_2\sigma^2_{PERSEflint}$. A covariance between the per se performance of an inbred and its corresponding GCA effect was considered in the model as follows. If GCA denotes a vector of GCA effects of all parents in one pool and $PERSE$ denotes a vector of the corresponding effects of line per se performance in the same pool and both vectors have the same order of genotypes, then the covariance of both vectors can be described by

$$\begin{aligned} \operatorname{cov}\begin{pmatrix} GCA \\ PERSE \end{pmatrix} &= \begin{pmatrix} \boldsymbol{A} \cdot \sigma^2_{GCA} & \boldsymbol{A} \cdot \sigma_{GCA;PERSE} \\ \boldsymbol{A} \cdot \sigma_{GCA;PERSE} & \boldsymbol{A} \cdot \sigma^2_{PERSE} \end{pmatrix} \\ &= \Sigma \otimes \boldsymbol{A}, \end{aligned}$$

where $\boldsymbol{A}$ is the additive relationship matrix, $\sigma_{GCA;PERSE}$ is the covariance between GCA and per se effect of parental lines, and Σ is a 2 × 2 unstructured matrix. This covariance structure was assumed for main GCA and line per se effects for both dent and flint lines as well as their interactions with location effects. All other interactions of genetic effects with year and location were assumed independent due to convergence issues, which may result from much stronger imbalance of parent-by-year classifications than of parent-by-location classifications. The residual error and the variance of incomplete blocks were assumed to be specific for each single trial. Since this model used pedigree (P+) data in the relationship matrices and considered covariance (C+) between GCA and line per se performance, the model was denoted as P+/C+ model. Three additional models were used, with varying variance–covariance structure. In the Model P+/C− all covariances between line per se performance and corresponding GCA effect were dropped from the model, so that Σ was diagonal. In Model P−/C+, identity matrices were used instead of A_1, A_2, and D. In Model P−/C− neither pedigree-based relationship matrices nor covariance between GCA and line per se performance were considered. These four models were compared on the basis of their restricted log-likelihood LL_R and the number of estimated parameters k by using Akaike's information criterion (AIC), which was calculated as $-2\ LL_R + 2k$, where models with lower AIC are preferable (Bozdogan 1987). All mixed linear model analyses were performed with ASReml (Gilmour et al. 2002). Coefficients of coancestry were calculated using the SAS procedure INBREED (SAS Institute Inc. 2000). The

P+/C+ model was chosen as the underlying model to study the influence on prediction efficiency of using (1) SSR instead of AFLP markers, (2) marker effects that were estimated separately for reciprocal allelic configurations of heterozygous marker loci, and (3) imputed missing AFLP marker data.

Molecular data

The 47 dent and 32 flint lines were genotyped with AFLP and SSR markers. The AFLP analyses were carried out with 20 primer-enzyme combinations (Vos et al. 1995), as described in detail by Schrag et al. (2006). Positions of 910 mapped AFLP bands were obtained from an integrated AFLP map (Vuylsteke et al. 1999). The SSR analyses were conducted with bulked samples of five plants per inbred line. Harvested leaves were freeze-dried and ground to powder. Genomic DNA was extracted using a modified CTAB procedure (Saghai-Maroof et al. 1984). Analyses were carried out for 256 publicly available SSR markers, which were uniformly distributed across the genome. Map positions (IBM2 2004 neighbors) and primer sequences were obtained from MaizeGDB (http://www.maizegdb.org). The primer pairs were synthesized by Sigma-Genosys (Steinheim, Germany), with one primer of each pair being fluorescent labeled with indodicarbocyanine (Cy5) at the 5′ end. The PCR reactions were performed in a total volume of 9.75 µl containing 60 ng of template DNA, 154 µM of each dNTP, 256 nM of each primer, 2.56 mM $MgCl_2$, 1 × PCR Buffer (Mg^{2+}-free), and 0.5 U *Taq* DNA polymerase (Invitrogen, Karlsruhe, Germany). Thermocycling consisted of an initial denaturation step of the template DNA at 94°C for 150 s, followed by 33 cycles of 93°C for 45 s, 52–60°C (depending on the primer set) for 45 s, and 72°C for 45 s, with a final extension phase of 10 min at 72°C. For some primer sets, a touchdown step from in between 58 and 64°C down to 55°C was included in the protocol. The resulting amplified DNA products were analyzed on polyacrylamide gels (Ultra Pure SequaGel-XR; National Diagnostics, Atlanta) on an ALF Express DNA sequencer (Amersham Biosciences, UK). The DNA fragments were sized with ALFWIN v2 software. Molecular markers were used for prediction if they were polymorphic and exhibited less than 30% missing observations, separately for dent and flint lines.

Imputation of missing AFLP marker data

Missing observations in the AFLP marker data set were replaced by predictions of marker genotypes, resulting in a complete marker data set, which was denoted as AFLP*. The missing genotypes were imputed based on measures of pair-wise haplotype dissimilarity within sliding windows of variable size, using NPUTE software (Roberts et al. 2007). Imputations were carried out separately for dent and flint lines and for each chromosome.

Cross-validation

Efficiency of prediction was evaluated by cross-validation with 300 randomized runs as described in detail by Schrag et al. (2009). For each cross-validation run, the 47 × 32 factorial data set was divided into an estimation set and a test set. In each cross-validation run, half of the parental lines from each heterotic group were randomly assigned as 'testcross-evaluated'. The crosses between testcross-evaluated lines formed the estimation set. The remaining crosses formed the test set, consisting of 'Type 0' crosses, meaning that none of the parents were evaluated for testcross performance, and 'Type 1' crosses, meaning that one parent was evaluated for testcross performance. In contrast to subsampling of testcross data, the per se performance data of all lines always remained in the estimation set. Variance components were estimated from the complete data set and used for all cross-validation runs. For each run, the Pearson correlation coefficient between predicted and observed performance was separately determined for Type 0 and Type 1 hybrids in the test set. The median across all runs of squared correlation coefficients (R^2) was considered as prediction efficiency. Boxplots of R^2 values were based on Tukey's five number summary (Tukey 1977).

Prediction methods

For HP prediction, six methods were studied, two of which were based on phenotypic and pedigree data (PP), and four on phenotypic, pedigree, and molecular marker data (Schrag et al. 2009). In the PP-GS method, HP was predicted by BLUPs for GCA and SCA of the corresponding dent and flint lines and their cross, estimated by mixed linear model analysis from the hybrids in the estimation set. If coefficients of coancestry were included in the mixed linear model analysis, the GCA estimates for testcross-unevaluated parental lines were obtained from data of related testcross-evaluated inbred lines, and the SCA estimates of untested hybrids were obtained from related tested hybrids. In the PP-L method, the average of BLUPs for per se performance of both parental lines was used as predictor of HP. Marker-based prediction methods were built on either multiple linear regression (MLR) or 'total effects of associated markers' ('TEAM'), both approaches being described in detail by Schrag et al. (2007). In short, for the MLR approach, performance of hybrids was regarded as a function of the effects of their allelic configurations at the marker loci in

the model. Variable selection of the marker loci was carried out by forward selection based on the F-to-enter procedure. For the TEAM approach, each marker locus was separately tested for significant association with hybrid performance at a false discovery rate of 5%. The effects of the allelic configurations on hybrid performance were then estimated and summed up across all significant loci, resulting in the TEAM value and considered as marker-based estimate of the genotypic value of a hybrid. In a simple linear regression across all hybrids, TEAM values were then used as predictors for hybrid performance (Schrag et al. 2007). In the MLR-H and TEAM-H methods, marker effects for HP were obtained from the estimation set and used for prediction of HP in the test set. In the analogous manner, marker effects were computed for mid-parent heterosis in the estimation set, and used for prediction of mid-parent heterosis in the test set. For the MLR-LM and TEAM-LM methods, these mid-parent heterosis predictions were combined with BLUPs of line per se performance, providing HP predictions of test set hybrids.

Distinguishing identical marker alleles from different heterotic groups

The marker genotype of each hybrid was determined by the marker genotypes of its parental inbred lines (Schrag et al. 2007). Given n marker alleles for a single locus, there are n^2 possible combinations of parental alleles. These consist of n homozygous and $n^2 - n$ heterozygous hybrid marker genotypes, the latter comprising $(n^2 - n)/2$ unique allele combinations and their $(n^2 - n)/2$ alternate counterparts. In a first approach (Het1), the reciprocal allelic configurations of heterozygous marker loci in hybrids were pooled together in the marker data as was done by Schrag et al. (2007, 2009). In a second approach (Het2), the reciprocal allelic configurations were distinguished for estimation of marker effects.

Results

Marker data

Assessment of the marker data for polymorphism and a maximum drop-out rate of 0.3 resulted in the retention of 732 markers in the AFLP data set, 891 markers in the AFLP* data set, and 179 markers in the SSR data set. The average numbers of allelic configurations at marker loci of hybrids for the AFLP data were 2.7 for Het1 and 3.5 for Het2, and these were very similar for AFLP*. The average numbers of allelic configurations for the SSR data set were 10.4 for Het1 and 12.2 for Het2.

Joint analyses of hybrids and inbreds

The number of estimated parameters k was 159 for P+/C+ and P−/C+ models, and 155 for P+/C− and P−/C− models. The models using pedigree-based relationship matrices (P+) showed lower AIC values (Table 1) than those using identity matrices (P−). Models including covariance between GCA and per se performance of the parental lines (C+) resulted in lower AIC than those, in which this covariance was not considered (C−). Consequently, the P+/C+ model showed the lowest AIC and the P−/C− model the highest. The estimates of variance components for GCA ranged from 0.175 to 0.226 for GY in Mg ha^{-1} and 1.89–2.92 for GDMC in percent (Table 1). Estimates for GCA and SCA variance components were higher for P+ models in most cases. The ratio of SCA:GCA variance components for the model with the lowest AIC (P+/C+ model) was 0.29 for GY and 0.14 for GDMC.

The differences in efficiencies between HP prediction approaches that were based on the four models P+/C+, P+/C−, P−/C+, and P−/C−, were largest for PP-GS and smallest for TEAM-H (Table 2). The prediction efficiencies were lower for MLR-H and MLR-PM (not shown) than for TEAM-H and TEAM-PM. However, the prediction efficiencies of the four models (P+/C+, P+/C−, P−/C+, P−/C−) showed similar patterns for MLR-H versus TEAM-H and for MLR-PM versus TEAM-PM. Use of the pedigree-based covariance matrices for GCA and SCA effects instead of identity matrices increased R^2 predominantly for PP-GS. With the P−/C− model, BLUP for GCA and SCA were zero due to the lack of covariance information between the genotypes, and, therefore, Type 0 hybrids were not predicted with PP-GS. However, while for GY these increases were large in C+ and C− models, for GDMC this was observed in C− models only. The inclusion of covariance between GCA and line per se performance in the mixed linear model enhanced R^2 predominantly for PP-GS, followed by PP-L and TEAM-LM, however, not for TEAM-H. This increase in R^2 of PP-GS was more pronounced for GDMC than for GY.

Modifications of the marker-based prediction methods

Markers associated with HP and mid-parent heterosis for GY and GDMC were selected with the marker-based prediction methods (Table 3). On the average across all runs, distinctly more markers were selected with TEAM than with MLR. When MLR methods were applied to SSR data, they selected only a very small number of markers and failed in some cases to select any marker. Distinguishing reciprocal allelic configurations of heterozygous marker loci in hybrids (Het2 vs. Het1) increased the number of

Table 1 Estimates of mixed linear model parameters from joint analyses of hybrids and parental lines for grain yield (in Mg ha^{-1}) and grain dry matter content (in %)

Trait/estimate	P+		P−	
	C+	C−	C+	C−
Grain yield				
LL_R	−28,114	−28,128	−28,152	−28,164
AIC	56,546	56,566	56,622	56,638
$\sigma^2_{GCAdent}$	0.226 ± 0.069	0.185 ± 0.067	0.175 ± 0.042	0.178 ± 0.042
$\sigma^2_{GCAflint}$	0.192 ± 0.074	0.197 ± 0.077	0.198 ± 0.064	0.200 ± 0.064
σ^2_{SCA}	0.061 ± 0.015	0.063 ± 0.016	0.043 ± 0.012	0.043 ± 0.012
Grain dry matter content				
LL_R	−8,394	−8,435	−8,491	−8,544
AIC	17,106	17,180	17,300	17,398
$\sigma^2_{GCAdent}$	2.92 ± 0.70	2.70 ± 0.67	2.44 ± 0.51	2.36 ± 0.49
$\sigma^2_{GCAflint}$	2.17 ± 0.66	2.21 ± 0.72	2.00 ± 0.59	1.89 ± 0.58
σ^2_{SCA}	0.36 ± 0.06	0.36 ± 0.06	0.28 ± 0.04	0.28 ± 0.04

Restricted log-likelihood (LL_R), Akaike's information criterion (AIC), and variance components (σ^2) of general (GCA) as well as specific combining ability (SCA) and their standard errors are given. The analyses were performed considering presence and absence of pedigree information (P+/P−) in combination with or without modelling the covariance between GCA and per se performance of parental lines (C+/C−)

selected markers for the TEAM methods, but reduced it for the MLR methods. For the imputed AFLP* data set, more markers were selected than for the original AFLP data set.

Prediction efficiencies for GY were highest for PP-GS (Table 4). For GY, the marker-based methods MLR-LM, TEAM-H and TEAM-LM were superior to PP-L, while for GDMC, only TEAM-LM resulted in similarly high R^2 as PP-L. The use of the SSR marker data set instead of AFLP for MLR-based prediction changed R^2 by −0.01 to +0.10 for Type 0 hybrids, and by −0.13 to +0.03 for Type 1 hybrids. With TEAM approaches, R^2 was increased by 0.03–0.04 for GY, but often reduced for GDMC (−0.05 to 0.02). Distinguishing identical marker alleles from different heterotic groups changed R^2 by −0.01 to +0.03 (data not shown). With the imputed AFLP* data set, R^2 of MLR-H was up to 0.06 higher than with the original AFLP data set. For the remaining marker-based approaches, R^2 changed by −0.02 to +0.02 when based on AFLP*.

Discussion

Joint analyses of hybrids and inbreds

Our study was based on field data that were also used by Schrag et al. (2009) for prediction of HP with mixed linear models. While performing the current analyses of the data we detected an incorrect ordering of columns and rows of the pedigree-based relationship matrices for GCA and line per se performance in the previous study. Owing to the corrected covariance matrices and some minor changes of the model in the present study, for P+/C+, the estimates of variance components and the R^2 of prediction based on GCA and line per se performance differed considerably from the earlier results in Schrag et al. (2009).

Variance components for GY in the present study were comparable to those reported by Fischer et al. (2008) for breeding material that covered a period of 30 years in the same breeding program. For GDMC, the variance components were in good agreement with Parisseaux and Bernardo (2004), who analyzed 22,774 single crosses belonging to nine heterotic patterns. Larger variance components for the P+ models in contrast to the P− models (Table 1) may be due to the reason that relationships among parental lines were considered by using pedigree-based covariance matrices instead of assuming unrelated parental lines by using identity matrices. To explain larger variance components for the P+ models, let us consider two genotypes with positive correlation b between them. Then, the $\boldsymbol{A}$-matrix can be written as $(1-b)\boldsymbol{I}+b\boldsymbol{J}$, where $\boldsymbol{I}$ is an identity matrix and $\boldsymbol{J}$ is a matrix with all entries equal to 1. In this case, $b\boldsymbol{J}$ is a constant effect for all genotypes and confounded with the overall mean of genotypes. For P− models, the $\boldsymbol{A}$-matrix is an identity matrix $\boldsymbol{I}$. For the considered case, $\boldsymbol{A}\boldsymbol{\sigma}^2$ is equal for P+ and P− models, from which follows, that $(1-b)\boldsymbol{I}\sigma^2_{p+}=\boldsymbol{I}\sigma^2_{p-}$, and thus for the estimated variance components, $\sigma^2_{p+}>\sigma^2_{p-}$.

Pedigree-based relationship matrices were used to model the covariance among inbred lines, both for their GCA and line per se performance. With these P+ models, R^2 for HP prediction of GY was greatly improved for PP-GS, but only marginally for PP-L and TEAM. These differences were due to considering the untested parental lines

Table 2 Efficiency of prediction (R^2) of grain yield (GY) and grain dry matter content (GDMC) of hybrids having no (Type 0) or one (Type 1) parental line evaluated in testcrosses

Trait/hybrid type	Method	Median R^2			
		P+		P−	
		C+	C−	C+	C−
Grain yield					
Type 1	PP-GS	0.73	0.61	0.53	0.39
	PP-L	0.32	0.14	0.23	0.13
	TEAM-H	0.54	0.54	0.50	0.50
	TEAM-LM	0.61	0.53	0.54	0.50
Type 0	PP-GS	0.60	0.40	0.22	NA
	PP-L	0.27	0.14	0.17	0.12
	TEAM-H	0.43	0.44	0.36	0.37
	TEAM-LM	0.46	0.39	0.37	0.35
Grain dry matter content					
Type 1	PP-GS	0.87	0.61	0.83	0.44
	PP-L	0.77	0.64	0.78	0.63
	TEAM-H	0.43	0.43	0.42	0.41
	TEAM-LM	0.80	0.72	0.80	0.72
Type 0	PP-GS	0.79	0.32	0.75	NA
	PP-L	0.75	0.64	0.75	0.63
	TEAM-H	0.23	0.24	0.22	0.23
	TEAM-LM	0.75	0.65	0.74	0.65

Prediction was based on phenotypic data (PP-GS, PP-L) or phenotypic and AFLP marker data (TEAM-H, TEAM-LM). The underlying joint analyses of hybrids and parental lines were performed considering presence and absence of pedigree information (P+/P−) in combination with or without modeling the covariance between general combining ability and per se performance of parental lines (C+/C−). The median prediction efficiency (R^2) was obtained across 300 cross-validation runs

NA not applicable

as having no testcrosses but being evaluated for line per se performance. Therefore, in our cross-validation approach, HP data were removed from the estimation sets, but data of line per se performance were retained. The GCA effects of parental lines that were not evaluated in testcrosses were not estimable from field data, but could be predicted from related genotypes by using pedigree-based covariance information. In contrast, the PP-L approach was based on the per se performance of lines, which was included in all estimation sets, resulting in only small additional benefit if pedigree-based covariance information was utilized. For GDMC there were practically no positive effects of P+ over P− under the C+ scenario, even for PP-GS. However, in the C− case, where PP-GS did not benefit from line per se GDMC performance data, additional information from related genotypes via pedigree-based relationship matrices improved the prediction efficiency of PP-GS. The improvement in R^2 of P+ was often larger for Type 0 than

Table 3 Average number of markers selected for prediction of hybrid performance and mid-parent heterosis of grain yield and grain dry matter content

Trait/marker data	Hybrid performance				Mid-parent heterosis			
	TEAM		MLR		TEAM		MLR	
	Het1	Het2	Het1	Het2	Het1	Het2	Het1	Het2
Grain yield								
AFLP	281.6	344.6	16.8	12.1	252.8	313.8	13.3	10.0
AFLP*	365.6	442.0	21.0	13.0	329.6	402.9	16.9	10.8
SSR	104.5	112.6	2.5	2.4	102.9	108.9	2.8	2.6
Grain dry matter content								
AFLP	233.3	296.6	20.6	13.9	141.8	189.9	14.7	10.7
AFLP*	300.3	381.1	26.0	14.1	181.0	242.7	19.7	12.4
SSR	92.1	101.0	2.5	2.2	76.1	82.4	2.1	2.0

Two methods (TEAM and MLR) were applied to three marker data sets, namely AFLP, AFLP* (AFLP with imputed missing observations), and SSR, to identify associated markers. Reciprocal allelic configurations of heterozygous marker loci in hybrids were pooled (Het1) or distinguished (Het2)

Table 4 Efficiency of prediction (R^2) for grain yield and grain dry matter content of hybrids having no (Type 0) or one (Type 1) parental line evaluated in testcrosses

Method	Marker data	Median R^2			
		Grain yield		Grain dry matter content	
		Type 1	Type 0	Type 1	Type 0
PP-GS	–	0.73	0.60	0.87	0.79
PP-L	–	0.32	0.27	0.77	0.75
MLR-H	AFLP	0.38	0.16	0.31	0.06
	AFLP*	0.44	0.21	0.31	0.09
	SSR	0.36	0.25	0.18	0.06
MLR-LM	AFLP	0.54	0.33	0.74	0.62
	AFLP*	0.54	0.33	0.71	0.60
	SSR	0.49	0.33	0.76	0.71
TEAM-H	AFLP	0.54	0.43	0.43	0.23
	AFLP*	0.54	0.45	0.43	0.25
	SSR	0.57	0.46	0.42	0.19
TEAM-LM	AFLP	0.61	0.46	0.80	0.75
	AFLP*	0.63	0.49	0.80	0.74
	SSR	0.65	0.50	0.82	0.76

Prediction was based on phenotypic data (PP-GS, PP-L) or phenotypic and AFLP marker data (MLR-H, MLR-LM, TEAM-H, TEAM-LM). The marker-based methods were applied to three marker data sets, namely AFLP, AFLP* (AFLP with imputed missing observations), and SSR. Reciprocal allelic configurations of heterozygous marker loci in hybrids were pooled for estimation of marker effects. The median prediction efficiency (R^2) was obtained across 300 cross-validation runs

for Type 1 hybrids, which was most evident with PP-GS. Since for Type 0 hybrids no testcross observations for the respective parental lines were available, the information gain due to the pedigree-based covariance matrix was more effective. For HP prediction, this demonstrated the advantage of covariance information among parents, if testcross information is lacking.

The correlation between the per se performance of parental inbred lines and their testcross performance is expected to be lower for heterotic traits, such as GY in maize, than for non-heterotic traits, such as GDMC (Hallauer and Miranda Filho 1988). Accordingly, for the C− models, the R^2 values of the PP-L approaches were among the lowest for GY and highest for GDMC. Consideration of covariance between GCA and per se performance of parental lines increased the efficiency of HP prediction for both traits and both PP-models, especially of GDMC for PP-GS. In the basic P−/C− model, the efficiency of the PP-GS method to predict GDMC of hybrids was much lower than for PP-L. Thus, for GDMC, PP-GS benefited from PP-L by taking into account the covariance between GCA and line per se performance, and in some cases even to such an extent, that R^2 with PP-GS were slightly higher than with PP-L. For Type 0 hybrids, no field data on testcrosses were available for estimation of GCA of the respective parental lines, so that pedigree-based relationship matrices and covariance between GCA and line per se performance were the only sources of information. This lack of other information may explain the fact that the beneficial effect of C+ for GDMC prediction with PP-GS was larger for Type 0 hybrids than for Type 1 hybrids and also larger, if no pedigree-based relationship matrices were used.

The use of pedigree-based relationship matrices and covariance between GCA and per se performance of parental lines enabled the exploitation of additional information from related genotypes and resulted in a better model fit to phenotypic data, indicated by the AIC values. In general, both enhancements also improved the efficiencies of HP prediction, yet to a varying degree, depending on the prediction method, trait and type of hybrid. The prediction approaches that were restricted to the use of phenotypic and pedigree data were more sensitive to the use of additional information from related genotypes than those based on markers (Fig. 1). For GY, the marker-based TEAM prediction approaches were often superior or at least similar to the PP approaches for their R^2, if no pedigree-based relationship matrices were used. In the P+/C− case, the TEAM approaches were comparable to PP-GS and clearly superior to PP-L. With the P+/C+ model, however, PP-GS achieved the highest R^2. For the prediction of GDMC based on C− models, the TEAM-L (marker-based) and PP-LM (not marker-based) methods were superior with respect to R^2. In summary, the results revealed the high predictive power of joint analyses of hybrids and lines, which utilized pedigree-based relationship matrices and covariance between GCA and per se performance of parental lines.

Enhancements of the marker-based prediction methods

Among the four mixed linear models, the P+/C+ model showed the lowest AIC for GY and GDMC, and therefore was chosen to further investigate the effects of modifications of the marker-based prediction methods on their efficiency. Efficiencies for prediction of HP using an AFLP marker data set were compared with that using an SSR marker data set, but did not reveal a general superiority of one over the other. Considering the use of mapped markers in this study, the total marker costs were estimated to be comparable or slightly lower for AFLP than for SSR data. The genome coverage in terms of the number of hybrid allele configurations across all loci was similar for AFLP and SSR. Due to higher costs and higher information

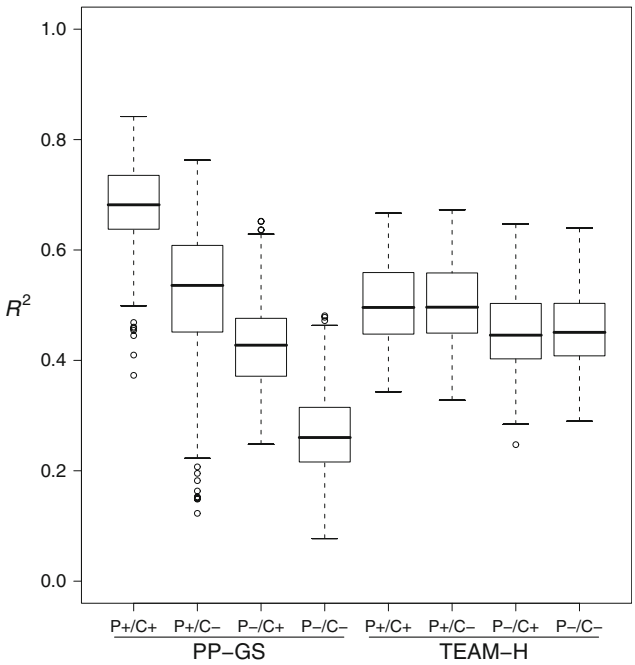

Fig. 1 Efficiency of prediction (R^2) of grain yield for hybrids having no or one parental line evaluated in testcrosses. Prediction was based on estimates of general and specific combining ability (PP-GS) or combining abilities and AFLP marker data (TEAM-H). The underlying joint analyses of hybrids and parental lines were performed considering presence or absence of pedigree information (P+/P−) in combination with or without modeling the covariance between general combining ability and per se performance of parental lines (C+/C−). Each of the eight *boxplot columns* corresponds to a specific prediction scenario and comprises 300 R^2 values, which in cross-validation were obtained from observed and predicted performance of hybrids

content per SSR marker locus, the initial number of available SSR markers was lower than for AFLP markers. Consequently, fewer markers were found to be significantly associated with the traits under study. Markers with missing observations in the respective data sample had to be excluded from the MLR analyses. This reduced the number of available SSR markers approximately by a factor of ten (data not shown), which can be the reason for the very low numbers of selected SSR markers for MLR.

When estimating the effects separately for reciprocal allelic configurations of heterozygous marker loci in hybrids, a reduced estimation error of marker effects and improved efficiency for HP prediction would be expected under the assumption that heterotic groups differ in their linkage disequilibria between marker and QTL alleles. The advantage of such a fine distinction of allelic configurations is associated with the disadvantage of higher numbers of configurations. Given n alleles for one locus, the maximum number of hybrid marker genotypes is $(n^2 + n)/2$ for Het1, and a higher number of n^2 configurations for Het2. Consequently, an increase of the average number of configurations was observed, resulting in a reduced number of observations per configuration. For the TEAM approaches, each marker was tested separately, so that the adverse effects on prediction may be smaller. However, in the MLR approach, all markers included in the model were fitted simultaneously. This could explain the considerable reduction in the number of selected markers. Both, advantage and disadvantage together, may explain the small differences in R^2 between Het1 and Het2. Population substructure within heterotic pools, e.g., Iodent and Stiff-Stalk subgroups within the dent pool, may be important. However, such further differentiation of alleles and their effects would aggravate the problem of large numbers of allelic configurations. Another problem when identifying markers associated with HP is the relationship among individuals. The true number of uncorrelated observations and therefore the degrees of freedom remain veiled, which may lead to false-positive marker associations in the MLR or TEAM approaches.

Missing marker observations in the AFLP data set were imputed using NPUTE software (Roberts et al. 2007). In this imputed data set AFLP* with no missing observations, the number of markers retained after assessment of data quality was higher (891 instead of 732). The number of markers that were selected by TEAM or MLR for their association with GY and GDMC was increased by a similar degree. However, R^2 values of TEAM differed only marginally between the imputation approach (AFLP*) and the original approach (AFLP), which substituted missing estimates of marker effects with average values. This indicated that both approaches were comparable in their potential to handle missing observations in the analyzed marker data set of AFLP markers. For the MLR approach, however, markers cannot be used for HP prediction if one or more observations are missing. Thus, improved R^2 were especially expected for the MLR approaches if applied to the AFLP* data set. This was observed for hybrid GY prediction with MLR-H only, however, not in the other cases. Altogether, the effects of imputing missing observations on HP prediction were small. With a drop-out rate of 1–3% and $N = 79$ genotyped inbreds, the expected proportion of usable markers is 9–45%, which is similar to what was observed for MLR in the current study (data not shown).

However, in commercial breeding programs with very large number of analyzed inbred lines, this limitation will be even more severe. With a low drop-out rate of 0.5% and with $N = 1{,}000$ genotyped lines, only 0.7% of all marker loci would be usable for MLR, which is a highly inefficient use of marker data. Thus, with an increasing amount of genotyping data, imputing missing marker observations seems to remain an issue. Also, for higher map densities, the efficiency of predicting missing marker observations from observed genotypes of tightly linked markers is expected to increase due to the higher linkage disequilibria. The joint effect of two modifications, namely (1) distinguishing reciprocal allelic configurations of heterozygous marker loci in hybrids (Het2) and (2) imputing missing observations in marker data (AFLP*) resulted in a slight increase of R^2 for prediction of GY mainly for MLR-H. In general, however, the modifications of the marker-based approaches had no clear effect (Fig. 2).

Conclusions for the application in breeding programs

For the prediction of GY and GDMC of hybrids, the PP-GS approach was superior to PP-L and marker-based approaches, if pedigree-based relationship measures and covariance between GCA and line per se performance were used. The choice of model for analysis of the phenotypic data had a considerably higher impact on R^2 than the modifications of marker-based approaches. It was demonstrated, that owing to the relatedness in the breeding materials and the relationship between hybrids and inbreds, the consideration of pedigree information (P+) and per se data (C+) resulted in the highest prediction efficiencies, especially for the non-marker approaches PP-GS and PP-L. For an untested hybrid, phenotypic data of its parents per se performance or of hybrids sharing a common parent can be regarded as estimators for the total effects of all trait-relevant genes in that hybrid. However, if no hybrids exist, that share one of the parents of the hybrid to be predicted, and also no per se performance of its parents is available, then marker-based approaches for prediction of HP have the potential to improve the efficiency of identifying superior hybrids. Molecular marker data may substitute

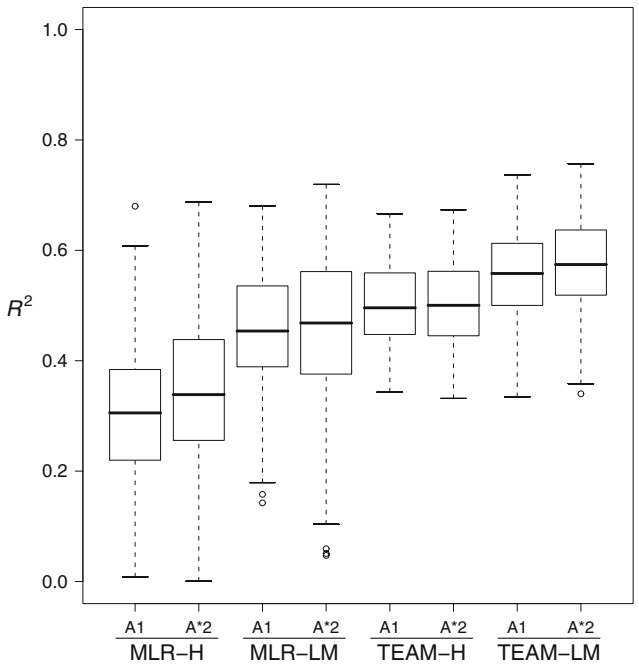

Fig. 2 Efficiency of prediction (R^2) of grain yield for hybrids having no or one parental line evaluated in testcrosses. The marker-based prediction approaches (MLR-H, MLR-LM, TEAM-H, TEAM-LM) were applied to AFLP data pooling the two alternate classes of each heterozygous hybrid marker genotype (A1) and to imputed AFLP* data distinguishing the two alternate classes of each heterozygous hybrid marker genotype (A*2). Each of the eight *boxplot columns* corresponds to a specific prediction scenario and comprises 300 R^2 values, which in cross-validation were obtained from observed and predicted performance of hybrids

pedigree data for the determination of the genotypic covariance matrix and could hereby further improve the prediction of hybrid performance.

Acknowledgments This article is dedicated to Prof. Dr. H. F. Utz on the occasion of his 70th birthday. The project was supported by the German Research Foundation DFG in the framework program 'Heterosis in Plants' (research grants FR 1615/3-1 and 3-3). We gratefully acknowledge the staff at the Plant Breeding Research Station at Eckartsweier and Hohenheim for conducting the field experiments and Christine Beuter for skilled technical assistance in the SSR marker lab. We thank Keygene N.V. for providing the AFLP marker data and mapping information. We thank the anonymous reviewers for their helpful comments and suggestions.

References

Balding DJ (2006) A tutorial on statistical methods for population association studies. Nat Rev Genet 7:781–791

Bernardo R (1992) Relationship between single-cross performance and molecular marker heterozygosity. Theor Appl Genet 83:628–634

Bernardo R (1994) Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci 34:20–25

Bernardo R (1999) Marker-assisted best linear unbiased prediction of single-cross performance. Crop Sci 39:1277–1282

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury

Bozdogan H (1987) Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika 52:345–370

Charcosset A, Essioux L (1994) The effect of population-structure on the relationship between heterosis and heterozygosity at marker loci. Theor Appl Genet 89:336–343

Charcosset A, Bonnisseau B, Touchebeuf O, Burstin J, Dubreuil P, Barriere Y, Gallais A, Denis JB (1998) Prediction of maize hybrid silage performance using marker data: comparison of several models for specific combining ability. Crop Sci 38:38–44

Dudley JW, Saghai Maroof MA, Rufener GK (1991) Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31:718–723

Fischer S, Möhring J, Schön CC, Piepho HP, Klein D, Schipprack W, Utz HF, Melchinger AE, Reif JC (2008) Trends in genetic variance components during 30 years of hybrid maize breeding at the University of Hohenheim. Plant Breed 127:446–451

Gardner CO, Eberhart SA (1966) Analysis and interpretation of the variety cross diallel and related populations. Biometrics 22:439–452

Gilmour AR, Cullis BR, Welham SJ, Thompson R (2002) ASReml Reference Manual. Release 1.0. VSN International, Hemel Hempstead

Hallauer AR, Miranda Filho JB (1988) Quantitative genetics in maize breeding. Iowa State University Press, Ames

Karakousis A, Barr AR, Chalmers KJ, Ablett GA, Holton TA, Henry RJ, Lim P, Langridge P (2003) Potential of SSR markers for plant breeding and variety identification in Australian barley germplasm. Aust J Agr Res 54:1197–1210

Lee M, Godshalk EB, Lamkey KR, Woodman WL (1989) Association of restriction length polymorphism among maize inbreds with agronomic performance of their crosses. Crop Sci 29:1067–1071

Maenhout S, De Baets B, Haesaert G (2010) Prediction of maize single-cross hybrid performance: support vector machine regression versus best linear prediction. Theor Appl Genet (this volume)

Parisseaux B, Bernardo R (2004) In silico mapping of quantitative trait loci in maize. Theor Appl Genet 109:508–514

Patterson HD (1997) Analysis of series of variety trials. In: Kempton RA, Fox PN (eds) Statistical methods for plant variety evaluation. Chapman & Hall, London, pp 139–161

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228

Rebourg C, Chastanet M, Gouesnard B, Welcker C, Dubreuil P, Charcosset A (2003) Maize introduction into Europe: the history reviewed in the light of molecular data. Theor Appl Genet 106:895–903

Roberts A, McMillan L, Wang W, Parker J, Rusyn I, Threadgill D (2007) Inferring missing genotypes in large SNP panels using fast nearest-neighbor searches over sliding windows. Bioinformatics 23:i401–i407

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

SAS Institute Inc (2000) SAS System 8. SAS Institute Inc, Cary

Schmidt W (2004) Hybridmaiszüchtung bei der KWS SAAT AG. Bericht über die 54. Tagung der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs 2003, Gumpenstein, pp 1–6 (in German)

Schrag TA, Melchinger AE, Sørensen AP, Frisch M (2006) Prediction of single-cross hybrid performance for grain yield and grain dry matter content in maize using AFLP markers associated with QTL. Theor Appl Genet 113:1037–1047

Schrag TA, Maurer HP, Melchinger AE, Piepho H-P, Peleman J, Frisch M (2007) Prediction of single-cross hybrid performance in maize using haplotype blocks associated with QTL for grain yield. Theor Appl Genet 114:1345–1355

Schrag TA, Möhring J, Maurer HP, Dhillon BS, Melchinger AE, Piepho H-P, Sørensen AP, Frisch M (2009) Molecular marker-based prediction of hybrid performance in maize using unbalanced data from multiple experiments with factorial crosses. Theor Appl Genet 118:741–751

Seitz G (2005) The use of doubled haploids in corn breeding. In: Proceedings of the 41st annual Illinois corn Breeders' School 2005, Urbana-Champaign, pp 1–7

Stich B, Melchinger AE, Frisch M, Maurer HP, Heckenberger M, Reif JC (2005) Linkage disequilibrium in European elite maize germplasm investigated with SSRs. Theor Appl Genet 111:723–730

Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP—a new technique for DNA-fingerprinting. Nucleic Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP (R) linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Vuylsteke M, Kuiper M, Stam P (2000) Chromosomal regions involved in hybrid performance and heterosis: their AFLP (R)-based identification and practical use in prediction models. Heredity 85:208–218