

# Performance prediction of F<sub>1</sub> hybrids between recombinant inbred lines derived from two elite maize inbred lines

Tingting Guo · Huihui Li · Jianbing Yan ·  
Jihua Tang · Jiansheng Li · Zhiwu Zhang ·  
Luyan Zhang · Jiankang Wang

Received: 20 December 2011 / Accepted: 16 August 2012  
© Springer-Verlag 2012

**Abstract** Selection of recombinant inbred lines (RILs) from elite hybrids is a key method in maize breeding especially in developing countries. The RILs are normally derived by repeated self-pollination and selection. In this study, we first investigated the accuracy of different models in predicting the performance of F<sub>1</sub> hybrids between RILs derived from two elite maize inbred lines Zong3 and 87-1, and then compared these models through simulation using a

wider range of genetic models. Results indicated that appropriate prediction models depended on genetic architecture, e.g., combined model using breeding value and genome-wide prediction (BV+GWP) has the highest prediction accuracy for high  $V_D/V_A$  ratio (>0.5) traits. Theoretical studies demonstrated that different components of genetic variance were captured by different prediction models, which in turn explained the accuracy of these models in predicting the F<sub>1</sub> hybrid performance. Based on genome-wide prediction model (GWP), 114 untested F<sub>1</sub> hybrids possibly having higher grain yield than the original F<sub>1</sub> hybrid Yuyu22 (the single cross between Zong3 and 87-1) have been identified and recommended for further field test.

Communicated by M. Frisch.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-012-1973-9) contains supplementary material, which is available to authorized users.

T. Guo · J. Li  
National Maize Improvement Center, China Agricultural University, Beijing 100193, China

T. Guo · H. Li · L. Zhang · J. Wang (✉)  
The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, No. 12 Zhongguancun South Street, Beijing 100081, China  
e-mail: jkwang@cgiar.org; wangjk@caas.net.cn

T. Guo · H. Li · L. Zhang · J. Wang  
CIMMYT China, Chinese Academy of Agricultural Sciences, Beijing, China

J. Yan  
National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

J. Tang  
Department of Agronomy, Henan Agricultural University, Zhengzhou 450002, China

Z. Zhang  
Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853, USA

## Introduction

Hybrid maize (*Zea mays* L.) breeders have been developing a large number of inbred lines and evaluating their performance in crosses (Hallauer 1990). The identification of elite F<sub>1</sub> hybrids between two inbred lines is the major objective in hybrid maize breeding. However, the number of potential crosses grows rapidly as more and more inbred lines are derived, and the field evaluation of hybrid performance requires large resources. For most hybrid breeding programs, only a small proportion of crosses can be evaluated in the field. An accurate prediction of hybrid performance prior to and after some field testing is of crucial importance in maize breeding.

A number of statistical models have been proposed to predict the hybrid performance in maize breeding. Hybrid prediction from inbred line per se performance is quite straight forward, but not effective because of the high level of dominance for grain yield. General combining ability (GCA) has been used in predicting hybrid performance

(Allard 1960; Cockerham 1967; Melchinger et al. 1987). GCA can be estimated by crossing individual inbred lines with as many other inbred lines as possible, along with various genetic mating designs including NCI (North Carolina design I), NCII, NCIII (Comstock and Robinson 1948, 1952), TTC (triple test crosses) (Kearsey and Jinks 1968) and Diallel design (Griffing 1956a, 1956b), etc. However, it ignores specific combining ability (SCA), which is related to heterosis and constitutes an important component of hybrid performance (Gardner and Eberhart 1966). Improving GCA-based model by taking SCA effects into consideration showed advantageous when SCA variances were of similar or larger importance compared with GCA variances (Schrag et al. 2006; Vuylsteke et al. 2000).

Genetic distance between parental lines based on random molecular markers, regarded as indicator of genetic diversity, can be used to predict hybrid performance. However, results from previous studies were inconsistent (Melchinger 1999; Jordan et al. 2003). Charcosset and Essioux (1994) attributed the low correlations between predicted and observed values to (1) no or only loose linkage of QTL affecting traits to molecular markers employed in estimating genetic distance and (2) different linkage phases between QTL and marker alleles exist in the maternal and paternal gametic arrays. Frisch et al. (2010) demonstrated that genetic distance based on transcription profiles was more precise than earlier prediction models using molecular markers.

Being a general method for predicting random effect, best linear unbiased prediction (BLUP) has been widely used in animal breeding (Henderson 1975, 1984), and was more recently advocated in plant breeding (Bernardo 1994, 1995, 1996a, 1996b, 1998, 1999). It is being used in genome-wide selection for quantitative traits in maize (Bernardo and Yu 2007). Recently BLUP has been compared to other statistical models for genome-wide selection such as support vector machine regression (Maenhout et al. 2010) and Bayesian methods (Meuwissen et al. 2001; Lorenzana and Bernardo 2009).

Mating and field experimental designs are crucial for genetic analysis and hybrid prediction. Hua et al. (2003) proposed a novel genetic population consisting of single crosses between a set of recombinant inbred lines (RILs) derived from two parental lines, which was called “immortalized  $F_2$ ” (or  $IF_2$ ). This population has the same genetic architecture as the conventional  $F_2$ , but each genotype in the population can be repeated and regenerated.  $IF_2$  population has been used to investigate the genetic basis of heterosis (Tang et al. 2010), but few studies have been conducted on performance prediction of  $F_1$  hybrid between RILs derived from two elite inbred lines in maize.

Hybrid prediction in bi-parental population is different from the prediction of inter-group hybrids for the following

reasons: (1) kinship between individuals in a bi-parental population is closer than those between two heterotic groups; (2) prediction in bi-parental population aims to improve the original  $F_1$  hybrid from two founders, through crossing their RILs. If a bi-parental RIL population is derived from a commercial hybrid between two elite inbred lines, it would be interesting to know whether there exist any  $F_1$  hybrids of two RILs resulting in better performance than the original  $F_1$  hybrid. Although the probability is low, and identified hybrid would make a significant impact. The objectives of this study were therefore (1) to construct statistical models to predict performance of  $F_1$  hybrids in a maize  $IF_2$  population; (2) to assess different prediction models in real and simulated populations, and (3) to identify potential RIL crosses outperforming the original  $F_1$  hybrid.

## Materials and methods

### Genetic population

Yuyu22 is an elite maize single cross in China and exhibits high level of heterosis on grain yield. One parent of Yuyu22 is the dent-inbred line Zong3 ( $P_1$ ), selected from a synthetic population of Chinese domestic germplasm, and the other parent is the flint-inbred line 871 ( $P_2$ ), selected from exotic germplasm. A total of 294 RILs were derived through single seed descent of Yuyu22. Similar to the procedure described by Hua et al. (2003), 294 RILs were randomly divided into two groups, each group having 147 RILs. Then, single crosses were randomly made between the two groups without replacement. This procedure was repeated three times, and finally 441 ( $147 \times 3$ ) single crosses were produced, forming the  $IF_2$  population.

Genotypic and phenotypic data used in this study were generated by the National Maize Improvement Center of China and was described in detail in Tang et al. (2010). The RIL and  $IF_2$  population were planted in 2003 and 2004 in Beijing (north of China, average daily temperature of 11.8 °C, and average annual rainfall of 585 mm) and Xunxian, Henan Province (center of the North China Plain, average daily temperature 14.2 °C, and average annual rainfall 784 mm). At two locations, populations of RIL and  $IF_2$  were in neighbored blocks, each planted in a randomized complete block design with three replications. Each plot included one row, 4-m long, with 0.67 m between rows. Population density was 45,000 plants per ha. After maturity, ten ears from consecutive plants in each plot were harvested by hand- and air-dried until the grain moisture reached 13 %. The 294 RILs and their 441  $F_1$  hybrids were evaluated for ten traits: ear weight (kg), grain yield (t/ha), kernels per row, ear diameter (cm), ear length (cm), row number, length of branch (cm), number of branches, ear

height (cm) and plant height (cm) (Table S1). The 294 RILs were screened by 261 polymorphic SSR markers covering the whole maize genome. Genotype of F<sub>1</sub> hybrid was deduced from its two RIL parents.

#### Analysis of phenotypic data

The phenotypic data were analyzed for sources of variation using the following model:

$$y_{ijk} = \mu + e_i + r_j + g_k + ge_{k \times i} + \varepsilon_{ijk}, \quad (1)$$

where  $y_{ijk}$  is trait value of the  $i$ th environment,  $j$ th replication and  $k$ th entry;  $\mu$  is the overall mean;  $e_i$  is the  $i$ th environment effect;  $r_j$  is the  $j$ th replication effect;  $g_k$  is the  $k$ th entry effect;  $ge_{k \times i}$  is the interaction effect of the  $k$ th entry by the  $i$ th environment;  $\varepsilon_{ijk}$  is the residual effect. In Eq. (1), all the variables except  $\mu$  are considered random effects. Broad sense heritability ( $H^2$ ) was equal to genetic variance ( $V_G$ ) divided by summation of genetic variance ( $V_G$ ), G  $\times$  E variance ( $V_{G \times E}$ ) and error variance ( $V_e$ ). Assuming the Yuyu22 F<sub>2</sub> is the reference population, genetic variance ( $V_G$ ) can be partitioned into additive and dominance variances ( $V_A$  and  $V_D$ ), ignoring the epistatic variance ( $V_I$ ). Three methods were used to dissect the two genetic variance components, and the best variance decomposition method was determined by simulation study. Firstly,  $V_A$  is included in genetic variance of the RIL population, and both  $V_A$  and  $V_D$  are included in genetic variance of the IF<sub>2</sub> population. So, under the condition that there are two equally frequent alleles in the population, we have,

$$V_A = \frac{1}{2}V_{G(\text{RIL})}, \text{ and } V_D = V_{G(\text{IF}_2)} - \frac{1}{2}V_{G(\text{RIL})}, \quad (2)$$

where  $V_{G(\text{RIL})}$  is genetic variance of the RIL population and  $V_{G(\text{IF}_2)}$  is genetic variance of the IF<sub>2</sub> population. Secondly, assuming no disequilibrium between loci, genetic variance can be estimated from QTL additive ( $a$ ) and dominance ( $d$ ) effects:

$$V_A = \frac{1}{2} \sum a^2, \text{ and } V_D = \frac{1}{4} \sum d^2. \quad (3)$$

Thirdly, assuming no disequilibrium between loci, genetic covariance between F<sub>1</sub> hybrids and their two parents is equal to the additive variance in bi-parental population (Table S2), which is different from the random-mating population where genetic covariance is half of additive variance. So additive and dominance variance can also be estimated by,

$$V_A = \text{cov}_{\text{OP}} = \text{cov} \left[ O, \frac{1}{2}(P_1 + P_2) \right], \text{ and} \quad (4)$$

$$V_D = V_{G(\text{IF}_2)} - V_A.$$

Narrow sense heritability ( $h^2$ ) was equal to additive variance ( $V_A$ ) divided by summation of genetic variance ( $V_G$ ), G  $\times$  E variance ( $V_{G \times E}$ ) and error variance ( $V_e$ ).

#### Prediction based on inbred line per se performance

Prediction based on inbred line per se performance was described in Hallauer et al. (2010), Smith (1986), and Zaidi et al. (2003). The predicted value of F<sub>1</sub> hybrid crossed by two RILs  $i$  and  $j$ , denoted by  $\hat{g}_{ij}$ , was calculated as,

$$\hat{g}_{ij} = \hat{\mu} + \frac{1}{2}(\hat{g}_i + \hat{g}_j), \quad (5)$$

where  $\hat{\mu}$  is the estimate of population mean from Eq. (1);  $\hat{g}_i$  and  $\hat{g}_j$  are genotypic values of the two parental lines, estimated by Eq. (1).

#### Prediction based on GCA

In IF<sub>2</sub> population, if several F<sub>1</sub> hybrids have one common parental line, GCA of this common parental line was calculated by means of its F<sub>1</sub> hybrids. GCA has been used to predict single, three-way and double crosses (Jenkins 1934; Cockerham 1967; Melchinger et al. 1987). GCA of the  $i$ th inbred line, defined as  $\text{GCA}_i$ , was estimated by:

$$\text{GCA}_i = \hat{\mu} + \frac{1}{2}(\hat{g}_{i \times m} + \hat{g}_{i \times n}), \quad (6)$$

where  $\hat{g}_{i \times m}$  and  $\hat{g}_{i \times n}$  are genotypic values of tested F<sub>1</sub> hybrids crossed by the  $i$ th inbred line with  $m$ th,  $n$ th inbred lines, estimated by Eq. (1). Then, the predicted value of F<sub>1</sub> hybrid was calculated by GCA of its two parental lines  $i$  and  $j$ :

$$\hat{g}_{ij} = \frac{1}{2}(\text{GCA}_i + \text{GCA}_j). \quad (7)$$

#### Prediction based on quantitative trait loci mapping

Combined QTL mapping was conducted in the mixture population of RILs and F<sub>1</sub> hybrids using inclusive composite interval mapping (Li et al. 2007; Zhang et al. 2008) implemented in the integrated software QTL IciMapping (<http://www.isbreeding.net>). The use of mixture population avoids the discrepancy of QTL location and effect when RIL and IF<sub>2</sub> populations are used separately. For the 1:1 mixture of RIL and IF<sub>2</sub>, three genotypes at one locus have frequencies 3/8, 1/4, and 3/8, similar to the genetic architecture of F<sub>3</sub> bulk populations. The observations in QTL mapping model are genotypic values of RILs and F<sub>1</sub> hybrids in mixture population. QTL mapping results used for prediction included (1) additive and dominance ( $a$  and  $d$ ) effects of QTL, and (2) genotypes of identified QTL of each RIL. Therefore, genotypes of QTL in untested F<sub>1</sub> hybrid were derived from its RIL parents, and predicted value of F<sub>1</sub> hybrid was calculated by

$$\hat{\mathbf{g}} = \hat{\mu} + \mathbf{M}\hat{\mathbf{a}} + \mathbf{N}\hat{\mathbf{d}}, \quad (8)$$

where  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{d}}$  are additive and dominance effects of QTL;  $\mathbf{M}$  and  $\mathbf{N}$  are incidence matrices where the number of rows

equals the size of predicted  $F_1$  hybrids and the number of columns equals the size of QTL. In matrix  $\mathbf{M}$ , 1, 0, and  $-1$  stands for QTL genotypes QQ, Qq, and qq, respectively; in matrix  $\mathbf{N}$ , 0 for homozygous QTL genotypes QQ and qq, and 1 for heterozygous QTL genotype Qq.

### Genome-wide prediction

Genotypic values and genotypes of RILs and their  $F_1$  hybrids were used to estimate  $a$  and  $d$  of molecular markers from a mixed model. The mixed linear model was denoted by

$$\hat{\mu} + \hat{g} = \mathbf{X}\beta + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{d} + \varepsilon, \quad (9)$$

where  $\hat{\mu} + \hat{g}$  is estimated by Eq. (1) and is the vector including values of RILs and  $F_1$  hybrids;  $\beta$  is the vector of fixed effect only including overall mean;  $\mathbf{X}$  is the incidence matrix with elements 1;  $\mathbf{a}$  and  $\mathbf{d}$  are vectors of additive and dominance effects, which are assumed to be random;  $\varepsilon$  is the residual vector;  $\mathbf{Z}_1$  is incidence matrix of  $\mathbf{a}$ , with elements 1,  $-1$ , and 0 corresponding to the three marker types MM, mm, and Mm, respectively;  $\mathbf{Z}_2$  is incidence matrix of  $\mathbf{d}$ , with elements 0 and 1 corresponding to the homozygous marker type and heterozygous marker type. The fixed and random effects were obtained by solving the mixed model equation;

$$\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{d}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_1 + \mathbf{A}^{-1} & \mathbf{Z}_1'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{Z}_2'\mathbf{R}^{-1}\mathbf{Z}_2 + \mathbf{D}^{-1} \end{bmatrix}^{-1} \times \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}(\hat{\mu} + \hat{g}) \\ \mathbf{Z}_1'\mathbf{R}^{-1}(\hat{\mu} + \hat{g}) \\ \mathbf{Z}_2'\mathbf{R}^{-1}(\hat{\mu} + \hat{g}) \end{bmatrix}, \quad (10)$$

$$\mathbf{A}^{-1} = \frac{m\sigma_\varepsilon^2}{\sigma_A^2}\mathbf{I}, \quad \text{and} \quad \mathbf{D}^{-1} = \frac{m\sigma_\varepsilon^2}{\sigma_D^2}\mathbf{I}$$

where residual matrix  $\mathbf{R}$  and  $\mathbf{I}$  are identity matrices;  $\sigma_A^2$ ,  $\sigma_D^2$  and  $\sigma_\varepsilon^2$  are additive variance, dominance variance and error variance, respectively;  $m$  is the number of markers. The genotypic value of  $F_1$  hybrid was predicted as:

$$\hat{g} = \mathbf{X}\hat{\beta} + \mathbf{Z}_1\hat{\mathbf{a}} + \mathbf{Z}_2\hat{\mathbf{d}}. \quad (11)$$

### Prediction based on breeding value

In the reduced animal model of BLUP (Quaas and Pollak 1980), breeding value of inbred line was calculated by:

$$\hat{\mu} + \hat{g} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \varepsilon, \quad (12)$$

where  $\hat{\mu} + \hat{g}$  was estimated by Eq. (1), representing genotypic values of RILs and  $F_1$  hybrids;  $\mathbf{u}$  is the vector of breeding values of RILs and  $\mathbf{Z}$  is the incidence matrix of  $\mathbf{u}$ .

Variable  $\mathbf{u}$  is assumed to be random and followed a multi-variation normal distribution,  $\mathbf{u} \sim \mathbf{N}(0, \mathbf{K}\sigma_u^2)$ , where the kinship matrix ( $\mathbf{K}$ ) was calculated using the marker data and the software SPAGeDi (Hardy and Vekemans 2002). Vectors  $\beta$  and  $\varepsilon$  and incidence matrix  $\mathbf{X}$  are defined as in Eq. (9).

### Prediction based on breeding value and genome-wide estimation of marker dominance effect

Dominance effect of molecular marker was estimated by a mixed model in  $IF_2$  population, in which the dependent value was the difference between  $\hat{\mu} + \hat{g}$  from Eq. (1) and breeding values  $\mathbf{u}$  from Eq. (12):

$$\hat{\mu} + \hat{g} - \mathbf{Z}\hat{\mathbf{u}} = \mathbf{X}\beta + \mathbf{Q}\mathbf{d} + \varepsilon, \quad (13)$$

where  $\mathbf{d}$  is the vector of dominance effect and  $\mathbf{Q}$  is incidence matrix of genotypes. The genotypic value of  $F_1$  hybrid was predicted as:

$$\hat{g} = \mathbf{X}\hat{\beta} + \mathbf{Z}\hat{\mathbf{u}} + \mathbf{P}\hat{\mathbf{d}}. \quad (14)$$

### Prediction in actual RIL and $IF_2$ populations in maize

All RILs together with the 294  $F_1$  hybrids ( $147 \times 2$ ) selected from the 441 hybrids ( $147 \times 3$ ) were used to build the prediction model, forming the training dataset. The remaining  $F_1$  hybrids were used to evaluate the effectiveness of prediction, forming the validation dataset. Performance of  $F_1$  hybrids in the validation dataset were estimated by the prediction model built in the training dataset, and squared correlation coefficient (represented by  $R^2$ ) between the predicted and observed performance was used to measure the accuracy of prediction. The procedure of sampling from 441 hybrids ( $147 \times 3$ ) was repeated three times, to give the standard error of  $R^2$ .

### Prediction in simulated populations

To assess and validate different prediction models, we conducted simulation studies on additive and dominance genetic model, and additive, dominance and epistatic genetic model. Linkage map built from the maize RIL population was used in simulation, and 50 QTL were assumed to control a trait in interest. In additive and dominance genetic model, we generated one hundred sets of random additive ( $a$ ) and dominance ( $d$ ) effects of the 50 QTL, where the genetic effects followed the uniform distribution between  $-1$  and 1. According to the effects of QTL, the ratio of dominance variance to additive variance ( $V_D/V_A$ ) was approximately calculated from  $\frac{1}{2} \sum_{m=1}^{50} d^2 / \sum_{m=1}^{50} a^2$ . Among the one hundred sets of QTL effects, four sets were chosen to represent four types of genetic

architectures, whose  $V_D/V_A$  was closest to 0, 0.5, 1, and 2, respectively (Table 1). When interaction between two QTL was considered in additive, dominance and epistatic genetic model, we also generated one hundred sets of random additive ( $a$ ) and dominance ( $d$ ) effects of 50 QTL and 10 pairs of epistatic effects ( $aa$ ,  $ad$ ,  $da$ , and  $dd$ ), where all effects followed the uniform distribution between  $-1$  and  $1$ . The  $V_I$  was approximately calculated by  $V_I = \sum (\frac{1}{4}aa^2 + \frac{1}{8}ad^2 + \frac{1}{8}da^2 + \frac{1}{16}dd^2)$ . Four of the one hundred sets of QTL effects were chosen to represent four types of genetic architectures, whose  $V_D/V_A$  was closest to 0.5, and  $V_I/(V_D + V_A)$  was closest to 0, 0.5, 1, and 2, respectively.

For each level of variance ratio, a total of 294 RILs were simulated by the genetics and breeding simulation tool of QuLine, formerly called QuCim (Wang et al. 2003, 2004), and 43,071  $F_1$  hybrids were made out of the 294 RILs. Similar to the actual populations, 294 RILs and 294  $F_1$  hybrids were used as the training dataset, and the other  $F_1$  hybrids were used as the validation dataset. The simulation was repeated three times, i.e. three sets of RIL and  $IF_2$  populations were constructed for each type of genetic architecture.

## Results

### Identification of significant QTL

In training dataset, the mixture population of RILs and  $IF_2$  had a size of 588. When the LOD threshold of 1.5 was applied, a total 187 QTL were identified for the ten traits (Table S3). They were distributed on the whole genome. Numbers of QTL ranged from 11 to 24, and total phenotypic variance explained by QTL was among 16.1 and 52.4 % across the ten traits. QTL identified in the mixture population were then used in prediction.

To show the advantage of using the mixture population in QTL mapping, we compared mapping results in the

mixture population with those in RIL and  $IF_2$  (Fig. S1). Under the LOD threshold of 1.5, less QTL were identified in the RIL population compared with  $IF_2$  and the mixture population. This is understandable because of (1) larger population size in  $IF_2$  and the mixture populations which have more power to detect QTL; (2) some QTL with significant dominance but non-significant additive effects. In the  $IF_2$  population, more QTL were detected for ear weight, grain yield, kernels per row and ear length. For other traits, the mixture population identified more QTL. Locations of most QTL from the three populations were overlapped. Additive effect estimated in the three populations was almost consistent, and dominance effect estimated in  $IF_2$  was greater than that in mixture populations (results not shown).

### Heritability, genetic variance and its components

Estimated heritability depends on which population is referred to. In the present study, broad sense heritability ( $H^2$ ) of each trait had different values when estimated from RIL,  $IF_2$  and the mixture populations (Fig. 1).  $H^2$  was the highest in the mixture population for ear weight, grain yield, kernels per row, ear diameter, ear length, row number, length of branch, plant height. For other two traits, number of branches and ear height,  $H^2$  was the highest in RILs.  $H^2$  in  $IF_2$  was the lowest for all traits except for ear weight and row number.

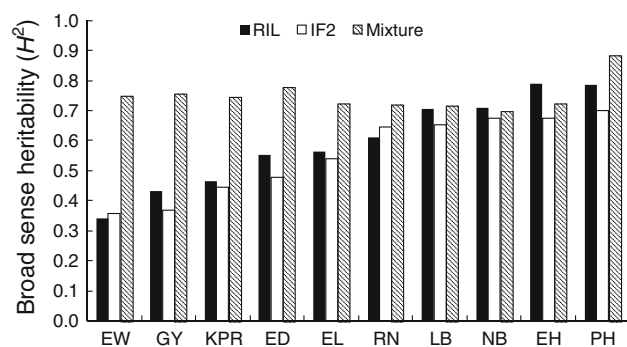
By ANOVA, phenotypic variance was divided into genetic, environmental,  $G \times E$ , and error variances for the ten traits in RIL,  $IF_2$  and the mixture populations (Table S4). Though having different values, variance components almost ranked in similar orders among the three kinds of populations. For all traits except ear weight in RILs, grain yield, and ear weight in  $IF_2$ , genetic variance had the highest value, indicating that genetic effect was the major factor on phenotypic performance of these traits.  $G \times E$  and environmental variances were also significant statistically on these traits.

**Table 1** Theoretical genetic variances in simulation study

| Genetic model                           | $V_A$ | $V_D$ | $V_I$ | $V_G$ |
|---|-------|-------|-------|-------|
| Additive and dominance model            |       |       |       |       |
| $V_D/V_A = 0.0$                         | 5.0   | 0.0   | 0.0   | 5.0   |
| $V_D/V_A = 0.5$                         | 5.0   | 2.5   | 0.0   | 7.5   |
| $V_D/V_A = 1.0$                         | 5.0   | 5.0   | 0.0   | 10.0  |
| $V_D/V_A = 2.0$                         | 5.0   | 10.0  | 0.0   | 15.0  |
| Additive, dominance and epistatic model |       |       |       |       |
| $V_I/(V_A + V_D) = 0.0, V_D/V_A = 0.5$  | 5.0   | 2.5   | 0.0   | 7.5   |
| $V_I/(V_A + V_D) = 0.5, V_D/V_A = 0.5$  | 5.0   | 2.5   | 3.7   | 11.2  |
| $V_I/(V_A + V_D) = 1.0, V_D/V_A = 0.5$  | 5.0   | 2.5   | 7.5   | 15.0  |
| $V_I/(V_A + V_D) = 2.0, V_D/V_A = 0.5$  | 5.0   | 2.5   | 15.0  | 22.5  |

$V_A$  additive variance,  $V_D$  dominance variance,  $V_I$  epistatic variance,  $V_G$  genetic variance





**Fig. 1** Broad sense heritability ( $H^2$ ) of ten traits calculated in the RIL, IF<sub>2</sub> and mixture populations. *EW*, *GY*, *KPR*, *ED*, *EL*, *RN*, *LB*, *NB*, *EH* and *PH* are abbreviations for ear weight, grain yield, kernels per row, ear diameter, ear length, row number, length of branch, number of branches, ear height, and plant height, respectively

Genetic variance was further dissected into additive and dominance components using three methods: ANOVA of RIL and IF<sub>2</sub> populations, summation of QTL *a* and *d* effects, and parent-offspring regression (Table 2). In the actual population, different methods resulted in different estimations of additive and dominance genetic variances.

For example,  $V_A$  and  $V_D$  of ear weight were estimated at 0.24 and 0.80 from ANOVA of RIL and IF<sub>2</sub> populations, 0.32 and 0.16 from summation of QTL *a* and *d* effects, and 0.14 and 0.90 from parent-offspring regression.

The best variance decomposition method can hardly be determined in actual populations, but can be properly determined in simulated populations, where true genetic variances are known. Under an additive and dominance model (Table 1), simulation studies demonstrated that  $V_A$  from ANOVA of RIL and IF<sub>2</sub> populations and from parent-offspring regression approximated their true values, e.g. the true  $V_A$  was 5.0, and  $V_A$  was estimated at 4.34 and 4.21–4.30, respectively (Table 2). However, estimated  $V_A$  from identified QTL effects was 4.76–11.90. The overestimation is likely caused by upward bias of effect estimates for detected, true and the presence of false positives resulting from the low LOD threshold applied for declaring statistical significance.  $V_D$  in different scenarios (Table 1) was slightly underestimated from ANOVA of RIL and IF<sub>2</sub> populations, and from parent-offspring regression, but was more clearly underestimated from identified QTL effects (Table 2). When epistatic effects were included in genetic

**Table 2** Additive and dominance variances estimated using three variance decomposition methods for ten traits in the actual and simulated populations

| Trait                     | $V_A$    |           |            | $V_D$    |           |            |
|---------------------------|----------|-----------|------------|----------|-----------|------------|
|                           | Method I | Method II | Method III | Method I | Method II | Method III |
| Actual population         |          |           |            |          |           |            |
| Ear weight                | 0.24     | 0.32      | 0.14       | 0.80     | 0.16      | 0.90       |
| Grain yield               | 0.27     | 0.28      | 0.09       | 0.51     | 0.04      | 0.69       |
| Kernels per row           | 5.27     | 1.94      | 1.54       | 2.62     | 0.69      | 6.35       |
| Ear diameter              | 0.05     | 0.03      | 0.03       | 0.00     | 0.01      | 0.02       |
| Ear length                | 1.27     | 1.27      | 0.64       | 0.82     | 0.17      | 1.46       |
| Row number                | 1.15     | 1.06      | 0.94       | 0.73     | 0.05      | 0.94       |
| Length of branch          | 6.13     | 3.76      | 4.93       | 0.98     | 0.24      | 2.17       |
| Number of branches        | 6.11     | 5.38      | 5.96       | 2.80     | 0.93      | 2.95       |
| Ear height                | 106.50   | 49.88     | 84.00      | 0.99     | 11.04     | 23.48      |
| Plant height              | 257.44   | 153.40    | 207.44     | 0.00     | 14.30     | 19.42      |
| Simulated population      |          |           |            |          |           |            |
| $V_D/V_A = 0.0$           | 4.34     | 11.90     | 4.30       | 0.00     | 0.07      | 0.10       |
| $V_D/V_A = 0.5$           | 4.34     | 8.23      | 4.26       | 2.19     | 2.45      | 2.26       |
| $V_D/V_A = 1.0$           | 4.34     | 5.75      | 4.21       | 4.64     | 3.61      | 4.77       |
| $V_D/V_A = 2.0$           | 4.34     | 4.76      | 4.22       | 8.94     | 7.39      | 9.05       |
| $V_I/(V_A + V_D) = 0.0^a$ | 4.34     | 8.23      | 4.26       | 2.19     | 2.45      | 2.26       |
| $V_I/(V_A + V_D) = 0.5^a$ | 5.71     | 6.28      | 5.25       | 4.33     | 2.38      | 4.79       |
| $V_I/(V_A + V_D) = 1.0^a$ | 9.07     | 10.89     | 6.45       | 4.69     | 4.16      | 7.31       |
| $V_I/(V_A + V_D) = 2.0^a$ | 10.32    | 10.16     | 8.06       | 11.17    | 4.91      | 13.42      |

Method I: variance estimated from ANOVA of RIL and IF<sub>2</sub> populations; Method II: variance estimated from summation of QTL *a* and *d* effects; Method III: variance estimated from parent-offspring regression

$V_A$  additive variance,  $V_D$  dominance variance

<sup>a</sup>  $V_I$  is epistatic variance, and the ratio of dominance variance and additive variance is equal to 0.5,  $V_D/V_A = 0.5$

model, i.e., under an additive, dominance and epistatic model (Table 1),  $V_A$  and  $V_D$  were both overestimated from all variance decomposition methods, compared with true  $V_A$  and  $V_D$  (Table 2). The overestimation might be explained by the inclusion of  $V_{AA}$  on the estimation of  $V_A$ , and the inclusion of  $V_{AD}$ ,  $V_{DA}$ , and  $V_{DD}$  on the estimation of  $V_D$ .

In summary, genetic variance decomposition using ANOVA of RIL and IF2 populations was slightly better than parent-offspring regression and much better than summation of QTL  $a$  and  $d$  effects. Even though the genetic variance components were biased estimated, rank of traits arranged by the ratio  $V_D/V_A$  was much the same for the three estimation methods. In the actual population,  $V_D/V_A$  of ear weight and grain yield was the highest but those of ear height and plant height were the lowest.

Correlation between predicted and observed  $F_1$  hybrid performances in the actual maize IF2 population

$R^2$  between predicted and observed performances of  $F_1$  hybrids varied among prediction models, e.g., for grain yield,  $R^2$  varied from 0.09 with ILP to 0.42 with BV+GWP (Table 3). In summary, QTL model had the lowest  $R^2$  (0.13–0.56); Model GCA had the intermediate  $R^2$  (0.27–0.53); ILP, GWP, BV and BV+GWP had the highest  $R^2$ , depending on the trait in consideration. Compared with other prediction models, ILP resulted in the highest  $R^2$  for number of branches (0.69), ear height (0.71) and plant height (0.66); GWP resulted in the highest  $R^2$  for ear diameter (0.56) and length of branch (0.65); BV resulted in the highest  $R^2$  for row number (0.66); BV+GWP resulted in the highest  $R^2$  for ear weight (0.44), grain yield (0.42), kernels per row (0.39) and ear length (0.49).

Prediction models ILP, GWP, BV and BV+GWP resulted in the highest  $R^2$  in at least one trait. Accuracy of these models in predicting  $F_1$  hybrid depended on the genetic architecture of the trait in interest (Fig. 2).  $H^2$  is the proportion of phenotypic variance explained by  $V_A$  and  $V_D$ , and  $h^2$  is that explained by  $V_A$ . In this sense, greater difference between  $H^2$  and  $h^2$  indicates more significance of  $V_D$ . In Fig. 2, the ten traits were arranged by  $H^2$  from the lowest to the highest. When  $H^2$  increased to 0.6, accuracy of the four models except for ILP almost reached their highest levels. For ear weight, grain yield, kernels per row, ear diameter, and ear length, when  $H^2 < 0.63$  and  $h^2 < 0.32$ , ILP had the poorest performance compared with other three models. On the contrary, BV+GWP had the best performance (Fig. 2). Furthermore, as the increase of heritability of traits with  $H^2 < 0.63$  and  $h^2 < 0.32$ , the performance of ILP in predicting  $F_1$  hybrids increased from 0.14 for ear weight to 0.42 for ear diameter. However, the performance of BV+GWP remains on high level

(0.39–0.54), showing stability of BV+GWP at different levels of heritability (Fig. 2). BV and GWP were also stable, but their performance could not exceed the performance of BV+GWP.

When  $H^2 > 0.63$  and  $h^2 > 0.32$  (Fig. 2), for length of branch, number of branches, ear height, and plant height, BV was the poorest prediction model except for its slightly better performance compared with BV+GWP for number of branches. ILP was the best, except for its slightly poorer performance compared with GWP for length of branch.  $H^2$  was the theoretical upper limit of  $R^2$  between observed and predicted  $F_1$  hybrids.  $R^2$  was equal to or greater than  $H^2$  in most traits (Fig. 2), indicating that the best prediction models have captured most, if not all, genetic variance. Any further improvement in prediction becomes less possible and more difficult. In some cases,  $R^2$  was greater than  $H^2$ , which could not be, but might be caused by sampling error, deviation in different models, or deviation when calculating  $H^2$ . For other traits, there might still have some space for further improvements.

Correlation between predicted and observed  $F_1$  hybrids in simulated populations

Prediction models were applied in simulated populations representing various genetic architectures. Results indicated that prediction accuracy of all models distinctly varied for traits with different  $V_D/V_A$ . When only additive effects were included in genetic model, all prediction models except GCA and QTL resulted in high  $R^2$  between predicted and observed  $F_1$  hybrids, more than 0.9 (Fig. 3). When  $V_D/V_A = 0.5$ ,  $V_D/V_A = 1$  and  $V_D/V_A = 2$ , the highest  $R^2$  were generated by GWP and BV+GWP. Model GWP performed as well as BV+GWP for traits with higher  $V_D/V_A$  ratios, which was inconsistent with what we found in actual maize population. The discrepancy was possibly caused by the more complicated genetic architecture in actual data.

In additive, dominance and epistatic genetic model, the trend of prediction model performance was the same as that in additive and dominance genetic model (Fig. 3). That was, BV+GWP and GWP were better than other prediction models as could be seen from the highest  $R^2$  between predicted and observed  $F_1$  hybrids. In comparison with additive and dominance genetic model, the  $R^2$  of all prediction models dropped off when epistatic variance existed in traits, and the prediction accuracy decreased along with the increase of  $V_I/(V_D + V_A)$ . Take GWP as an example, when  $V_I/(V_D + V_A)$  was increased from 0 to 2,  $R^2$  decreased from 0.78 to 0.40. Therefore, the presence of epistatic variance complicates all prediction models and reduces the prediction accuracy. High-accuracy prediction can only be achieved when  $V_A$  and  $V_D$  are the major genetic variance.

**Table 3** Squared correlation coefficients ( $R^2$ ) between predicted and observed  $F_1$  hybrid performance for ten traits using 7 prediction models

| Model               | Ear weight               | Grain yield | Kernels per row | Ear diameter | Ear length  | Row number  | Length of branch | Number of branches | Ear height  | Plant height |
|---------------------|--------------------------|-------------|-----------------|--------------|-------------|-------------|------------------|--------------------|-------------|--------------|
| ILP <sup>a</sup>    | 0.14 ± 0.02 <sup>g</sup> | 0.09 ± 0.06 | 0.14 ± 0.11     | 0.42 ± 0.01  | 0.24 ± 0.04 | 0.60 ± 0.02 | 0.61 ± 0.01      | 0.69 ± 0.04        | 0.71 ± 0.02 | 0.66 ± 0.03  |
| GCA <sup>b</sup>    | 0.28 ± 0.04              | 0.27 ± 0.05 | 0.31 ± 0.04     | 0.43 ± 0.02  | 0.35 ± 0.03 | 0.53 ± 0.04 | 0.51 ± 0.04      | 0.45 ± 0.05        | 0.45 ± 0.03 | 0.46 ± 0.05  |
| QTL <sup>c</sup>    | 0.13 ± 0.08              | 0.13 ± 0.09 | 0.15 ± 0.05     | 0.35 ± 0.04  | 0.24 ± 0.07 | 0.56 ± 0.03 | 0.49 ± 0.04      | 0.45 ± 0.04        | 0.51 ± 0.04 | 0.45 ± 0.06  |
| GWP <sup>d</sup>    | 0.36 ± 0.01              | 0.35 ± 0.02 | 0.32 ± 0.01     | 0.56 ± 0.02  | 0.48 ± 0.02 | 0.63 ± 0.02 | 0.65 ± 0.02      | 0.58 ± 0.02        | 0.61 ± 0.05 | 0.64 ± 0.04  |
| BV <sup>e</sup>     | 0.34 ± 0.02              | 0.31 ± 0.02 | 0.30 ± 0.05     | 0.50 ± 0.02  | 0.40 ± 0.02 | 0.66 ± 0.03 | 0.58 ± 0.02      | 0.56 ± 0.02        | 0.42 ± 0.04 | 0.47 ± 0.01  |
| BV+GWP <sup>f</sup> | 0.44 ± 0.00              | 0.42 ± 0.02 | 0.39 ± 0.03     | 0.54 ± 0.02  | 0.49 ± 0.02 | 0.65 ± 0.02 | 0.61 ± 0.02      | 0.55 ± 0.02        | 0.44 ± 0.05 | 0.50 ± 0.02  |

Data in italic font show the value of  $R^2$  is the largest among all prediction models for each trait

<sup>a</sup> Prediction based on inbred line per se performance

<sup>b</sup> Prediction based on general combining ability (GCA)

<sup>c</sup> Prediction based on QTL mapping

<sup>d</sup> Genome-wide prediction

<sup>e</sup> Prediction based on breeding value

<sup>f</sup> Prediction based on breeding value and genome-wide estimation of marker dominance effect

<sup>g</sup> The number following by “±” is standard error of  $R^2$  through three times sampling of training set

Prediction of elite  $F_1$  hybrids and elite recombinant inbred lines

Prediction models ILP, GWP, BV and BV+GWP produced the highest  $R^2$  in corresponding traits, and were used to illustrate the prediction of untested  $F_1$  hybrids. Predicted values by the four models were different even for the same  $F_1$  hybrid, due to some systematic errors. The best 50 combinations based on prediction were shown in Table S5, from which the best  $F_1$  hybrid and RIL could be determined by predicted performance and the repeatability in the four prediction models. Taking grain yield as an example, hybrid RIL90 × RIL153 ranked the first in prediction models BV and BV+GWP, and the second for GWP. RIL185, RIL153, RIL187 and RIL90 repeatedly appeared 41, 32, 30 and 29 times in the first 50 best  $F_1$  hybrids out of all untested combinations (Table S5). Therefore, the  $F_1$  hybrid RIL90 × RIL153 and inbred lines RIL185, RIL153, RIL187 and RIL90 could be considered as the best RIL crosses and lines for further field test and breeding.

Genotypic values of the 42,630 untested  $F_1$  hybrids were predicted using GWP, showing satisfactory accuracy in both actual and simulated populations. Maximum genotypic values across the untested  $F_1$  hybrids were greater than those of the original  $F_1$  for all traits (Table 4). For example, ranges of genotypic values of all untested  $F_1$  hybrids were from 5.20 to 7.90 for grain yield, and the original  $F_1$  hybrid was estimated at 7.60 for grain yield. For plant height and ear height, it should be noted that lower values may be favored and minimum genotypic values across the untested  $F_1$  hybrids were lower than those of the original  $F_1$  (Table 4).

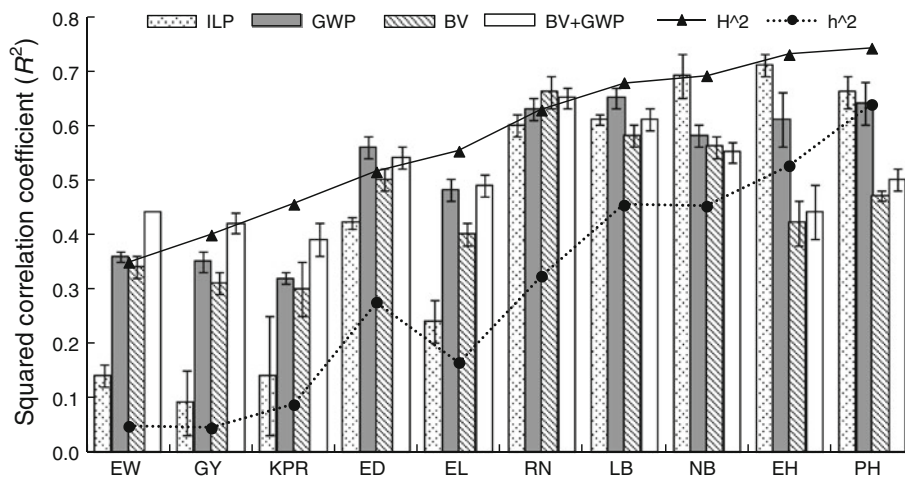
Fewer untested  $F_1$  hybrids were identified to be superior to the original  $F_1$  for ear weight and grain yield (Table 4). The proportions were so low (0.08 % for ear weight; 0.27 % for grain yield) that it would be very difficult to select the superior ones without efficient prediction. For other traits, especially ear diameter and ear height, it might be easy to select  $F_1$  hybrids with higher or lower genotypic values, because the proportions of superior  $F_1$  hybrids were high enough, e.g., 78.29 % for ear diameter and 74.85 % for ear height. Among the tested  $F_1$  hybrids (147 × 3), one hybrid was identified to be superior to the original  $F_1$  in ear weight, four hybrids superior to the original  $F_1$  in grain yield, and eight hybrids superior to the original  $F_1$  in kernels per row (Table 4). For other traits, more tested  $F_1$  hybrids were identified to be superior to the original.

## Discussion

Genetic architecture determines accuracy of prediction

Hybrid prediction can help breeders identify superior  $F_1$  hybrids and reduce the number of hybrids for field testing.

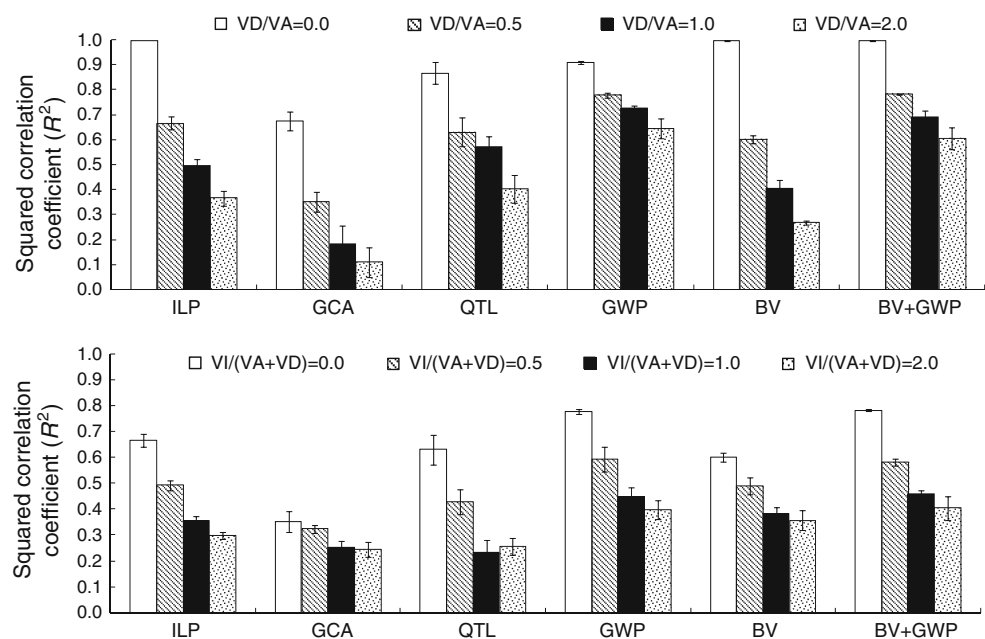




**Fig. 2** Squared correlation coefficients ( $R^2$ ) between observed and predicted  $F_1$  hybrid performance of prediction models ILP, GWP, BV, and BV+GWP in the actual population. The *solid line* and *dotted line* are broad sense heritability ( $H^2$ ) and narrow sense heritability ( $h^2$ ) for ten traits, respectively. *EW*, *GY*, *KPR*, *ED*, *EL*, *RN*, *LB*, *NB*, *EH* and *PH* are abbreviations for ear weight, grain yield, kernels per row, ear

diameter, ear length, row number, length of branch, number of branches, ear height and plant height, respectively. ILP is based on inbred line per se performance; GWP is genome-wide prediction; BV is based on breeding value; BV+GWP based on breeding value and genome-wide estimation of marker dominance effect

**Fig. 3** Squared correlation coefficients ( $R^2$ ) between observed and predicted  $F_1$  hybrid performance of prediction models in simulated population for additive and dominance genetic model (*upper*), and additive, dominance, and epistatic genetic model (*lower*). ILP is based on inbred line per se performance; GCA based on general combining ability; QTL based on QTL mapping; GWP is genome-wide prediction; BV based on breeding value; BV+GWP based on breeding value and genome-wide estimation of marker dominance effect.  $V_A$ ,  $V_D$ ,  $V_I$  are additive, dominance, and epistatic variance, respectively



New available information and tools on molecular markers, advanced genetic designs and statistical methods have accelerated the development of new prediction models in recent years (Bernardo 1995; Meuwissen et al. 2001; Jordan et al. 2003; Schrag et al. 2006). In this study, we evaluated six prediction models in an immortal  $F_2$  population and in simulated populations. Theoretical interpretations were given on the observed prediction effectiveness.

Assuming a trait of interest is controlled by two genes, denoted by  $A/a$  and  $B/b$  for demonstration, there are four genotypes in the RIL population and ten genotypes in the

$IF_2$  population by considering the coupling and repulsive linkage phases. Assuming no segregation distortion in RIL population, each allele frequency is 0.5. If two loci are independent, the theoretical genotypic frequency of each genotype should be 0.25 (Table S6B). Theoretical genotypic values of RILs and  $F_1$  hybrids in  $IF_2$  could be obtained from their genotypes (Table S6B and Table 5). The predicted genotypic values for models BV, ILP, GWP and BV+GWP were shown in Table 5. For model BV (Table 5), predicted value of  $F_1$  hybrid was the mean value of breeding values of its parents, and breeding value was

equal to the sum of average effects of the genes it carried (Tables S6A, S6B). Theoretical genotypic values of  $F_1$  hybrids included additive, dominance and epistatic effects; in contrast, predicted values included only genetic effects related to additive effects ( $a_A$ ,  $a_B$ ,  $aa_{AB}$ ,  $ad_{AB}$  and  $da_{AB}$ ). Therefore, the difference between theoretical and predicted values contained dominance effect and partial epistasis. When model ILP was used for  $F_1$  hybrid prediction (Table 5), the predicted value was the average of parental theoretical genotypic values which were shown in Table S6B. The difference between theoretical and predicted values was 0 if the offspring was homozygous. Otherwise, dominance and all types of epistatic effects remained in the difference between theoretical and predicted values. In other words, genetic effects related to dominance ( $d_A$ ,  $d_B$ ,  $ad_{AB}$ ,  $da_{AB}$  and  $dd_{AB}$ ) could not be captured by ILP.

Additive and dominance effects of all loci could be predicted by GWP using Eq. (9), then, predicted value of  $F_1$  hybrid included all additive and dominance effects (Table 5), and there was no difference between theoretical and predicted values under additive and dominance genetic model. However, under additive, dominance, and epistatic genetic model, GWP could not capture any type of epistasis, therefore, the difference between theoretical and predicted values included all types of epistasis. Under prediction model BV+GWP (Table 5), breeding value of  $F_1$  hybrid was equal to the sum of average effects of the genes it carried (Tables S6A and S6B), and dominance effects of all loci were predicted using Eq. (13). Thus, all additive and dominance effects, and half epistasis in theoretical genotypic values could be captured in the predicted genotypic values (Table 5). Among these prediction models, additive effect could be captured by all models, dominance effect could be captured by GWP and BV+GWP, and epistasis could be captured by BV (partial), ILP ( $aa_{AB}$ )

and BV+GWP. Therefore, BV+GWP was most efficient as it captured all types of genetic effects.

$F_1$  hybrid of two RILs consists of a range of homogeneous and heterogeneous genotypes. The genetic effects in  $F_1$  hybrid include additive, dominance, and epistatic effects, which determine the accuracy of prediction. For traits mainly controlled by additive genes, most prediction models could achieve satisfactory accuracy. For traits with a complicated genetic architecture, the six prediction models in this study had different accuracies (Table 5; Fig. 2). BV+GWP had the ability to capture most genetic variation, resulting in the highest accuracy among the prediction models considered in this study. It should be noted that the best theoretical model may not perform well in every practical dataset, due to biased genetic effects.

#### Efficiency of different prediction models

In the present study, phenotypic data, molecular marker data and genetic relationship between individuals were the required input for predicting hybrid performance. Among the six models, ILP and GCA only utilized phenotype, and their prediction accuracy was lower than those of GWP, BV, and BV+GWP for most traits, including ear weight, grain yield, kernels per row, ear diameter, ear length, and row number. Molecular markers help to dissect the genetic architecture of breeding traits and increase the accuracy of prediction. Therefore, in the majority of cases, GWP was better than ILP and GCA which used only phenotype of RILs and  $F_1$  hybrids. The use of genetic relationship between individuals increased the prediction accuracy as well, as can be seen from the higher accuracy of BV, and BV+GWP compared with ILP, QTL, and GWP. These observations are consistent with the results shown by De Los Campos et al. (2009). In conclusion, the more genetic

**Table 4** Predicted genotypic values of tested and untested  $F_1$  hybrids in the actual population using genome-wide prediction (GWP) model

| Trait              | Genotypic value of original $F_1$ hybrid | Range of genotypic values of untested $F_1$ hybrids | Number of untested superior $F_1$ hybrids <sup>a</sup> | Proportion of untested superior $F_1$ hybrids (%) | Range of genotypic values of tested $F_1$ hybrids | Number of tested superior $F_1$ hybrids | Proportion of tested superior $F_1$ hybrids (%) |
|--------------------|--|---|--|---|---|---|---|
| Ear weight         | 8.99                                     | 6.36–9.21   | 33   | 0.077   | 6.77–8.99   | 1                                       | 0.227   |
| Grain yield        | 7.60                                     | 5.20–7.90   | 114  | 0.265   | 5.72–7.76   | 4                                       | 0.907   |
| Kernels per row    | 33.16                                    | 25.18–34.75   | 524  | 1.217   | 26.96–34.25                                       | 8                                       | 1.814   |
| Ear diameter       | 4.67                                     | 4.22–5.36   | 33,720   | 78.289  | 4.38–5.16   | 343                                     | 77.778  |
| Ear length         | 17.25                                    | 13.56–19.17   | 6,815  | 15.823  | 14.33–18.55                                       | 76                                      | 17.234  |
| Row number         | 16.40                                    | 13.03–19.62   | 9,649  | 22.403  | 13.43–18.54                                       | 101                                     | 22.902  |
| Length of branch   | 35.06                                    | 26.96–40.11   | 13,699   | 31.806  | 28.32–39.29                                       | 143                                     | 32.426  |
| Number of branches | 16.27                                    | 9.18–21.73  | 12,019   | 27.905  | 10.12–20.55                                       | 136                                     | 30.839  |
| Ear height         | 101.78                                   | 64.11–125.97  | 31,909   | 74.851  | 75.43–116.29                                      | 318                                     | 72.109  |
| Plant height       | 220.33                                   | 185.01–271.08                                       | 7,055  | 16.549  | 201.64–262.86                                     | 72                                      | 16.327  |

<sup>a</sup> The superior  $F_1$  hybrids are the ones whose genotypic values are higher (or lower for plant height and ear height) than original  $F_1$  hybrid

**Table 5** Theoretical genotypic values and predicted values of the ten possible genotypes in IF<sub>2</sub> when the trait is controlled by two genes

| Genotypes          | Theoretical value                  | Predicted value   |                        |                  |   |
|--------------------|------------------------------------|---|------------------------|------------------|---|
|                    |                                    | BV <sup>a</sup>   | ILP <sup>b</sup>       | GWP <sup>c</sup> | BV + GWP <sup>d</sup>   |
| AABB               | $a_A + a_B + aa_{AB}$ <sup>e</sup> | $a_A + a_B + \frac{1}{2}aa_{AB} + \frac{1}{2}ad_{AB} + \frac{1}{2}da_{AB}$  | $a_A + a_B + aa_{AB}$  | $a_A + a_B$      | $a_A + a_B + \frac{1}{2}aa_{AB} + \frac{1}{2}ad_{AB} + \frac{1}{2}da_{AB}$  |
| AABb               | $a_A + d_B + ad_{AB}$              | $a_A + \frac{1}{2}ad_{AB}$  | $a_A$                  | $a_A + d_B$      | $a_A + d_B + \frac{1}{2}ad_{AB}$  |
| AaBB               | $d_A + a_B + da_{AB}$              | $a_B + \frac{1}{2}da_{AB}$  | $a_B$                  | $d_A + a_B$      | $a_B + d_A + \frac{1}{2}da_{AB}$  |
| AB/ab <sup>f</sup> | $d_A + d_B + dd_{AB}$              | $\frac{1}{2}aa_{AB}$  | $aa_{AB}$              | $d_A + d_B$      | $d_A + d_B + \frac{1}{2}aa_{AB}$  |
| AAbb               | $a_A - a_B - aa_{AB}$              | $a_A - a_B - \frac{1}{2}aa_{AB} + \frac{1}{2}ad_{AB} - \frac{1}{2}da_{AB}$  | $a_A - a_B - aa_{AB}$  | $a_A - a_B$      | $a_A - a_B - \frac{1}{2}aa_{AB} + \frac{1}{2}ad_{AB} - \frac{1}{2}da_{AB}$  |
| Ab/aB <sup>g</sup> | $d_A + d_B + dd_{AB}$              | $-\frac{1}{2}aa_{AB}$   | $-aa_{AB}$             | $d_A + d_B$      | $d_A + d_B - \frac{1}{2}aa_{AB}$  |
| Aabb               | $d_A - a_B - da_{AB}$              | $-a_B - \frac{1}{2}da_{AB}$   | $-a_B$                 | $d_A - a_B$      | $-a_B + d_A - \frac{1}{2}da_{AB}$   |
| aaBB               | $-a_A + a_B - aa_{AB}$             | $-a_A + a_B - \frac{1}{2}aa_{AB} - \frac{1}{2}ad_{AB} + \frac{1}{2}da_{AB}$ | $-a_A + a_B - aa_{AB}$ | $-a_A + a_B$     | $-a_A + a_B - \frac{1}{2}aa_{AB} - \frac{1}{2}ad_{AB} + \frac{1}{2}da_{AB}$ |
| aaBb               | $-a_A + d_B - ad_{AB}$             | $-a_A - \frac{1}{2}ad_{AB}$   | $-a_A$                 | $-a_A + d_B$     | $-a_A + d_B - \frac{1}{2}ad_{AB}$   |
| aabb               | $-a_A - a_B + aa_{AB}$             | $-a_A - a_B + \frac{1}{2}aa_{AB} - \frac{1}{2}ad_{AB} - \frac{1}{2}da_{AB}$ | $-a_A - a_B + aa_{AB}$ | $-a_A - a_B$     | $-a_A - a_B + \frac{1}{2}aa_{AB} - \frac{1}{2}ad_{AB} - \frac{1}{2}da_{AB}$ |

<sup>a</sup> Prediction based on breeding value, and predicted value is the mean values of breeding values of its parents which is shown in Table S6B

<sup>b</sup> Prediction based on inbred line per se performance, and predicted value is the average of parental genotypic values which is shown in Table S6B

<sup>c</sup> Genome-wide prediction

<sup>d</sup> Prediction based on breeding value and genome-wide estimation of marker dominance effect

<sup>e</sup>  $a$  and  $d$  are additive and dominance effects;  $aa$ ,  $ad$ ,  $da$ ,  $dd$  are four types of epistatic effects

<sup>f</sup> AABB × aabb

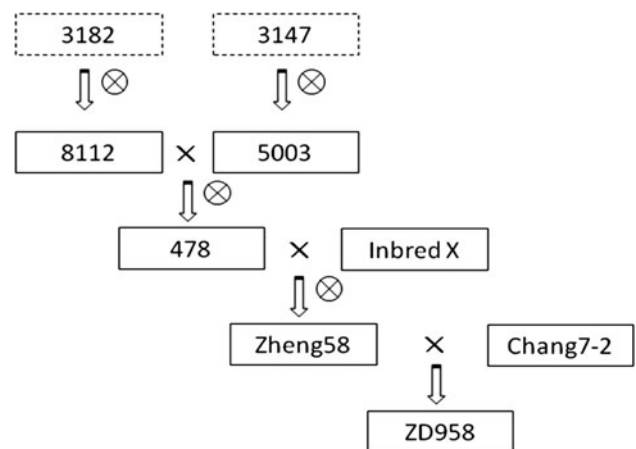
<sup>g</sup> AAaa × aaBB

information used, the better and stable prediction model performed. It is anticipated that the accuracy from BV+GWP could be further improved when more epistatic effects can be included in prediction. Statistical method needs to be improved so as to catch epistatic effects in the linear prediction model in GWP where the variables representing epistasis greatly exceed the population size.

It was clearly shown that the best prediction model depends on the genetic architecture of the trait in interest. The information regarding which model is the best can be helpful if the genetic architecture can be identified from additional genetic data. Grain yield is of primary importance, but other traits such as grain moisture at harvest, lodging resistance, grain quality are also important (Hallauer et al. 2010). The use of multiple traits gave us the chance to see that the best prediction model depends on the genetic architecture of the trait in interest.

#### Applications of prediction in maize breeding

Maize breeding involves two critical steps, developing superior inbred lines from breeding populations and identifying elite combinations of two inbred lines. In US, inbred lines are commonly developed from F<sub>2</sub> populations made from elite × elite F<sub>1</sub> hybrids of related lines within heterotic groups (Hallauer et al. 2010). In developing countries, like China, however, maize breeding is conducted by many small public sectors. The heterotic groups have not been clearly defined, and inbred lines are



**Fig. 4** Pedigree of the Chinese maize hybrid ZD958 (adapted from Lai et al., 2010). The female parent of each cross is listed first. Zheng58's other parent is unknown and termed as 'Inbred X'. The US commercial hybrids are framed by dotted line

commonly developed from the segregating populations derived from commercial hybrids (Yu et al. 2007). For example, ZD958 is an F<sub>1</sub> hybrid released in 2000 in China. The planting area reached 3.906 million ha in 2006 and is currently the most widely grown in China. Zheng58, one parental inbred line of ZD958, was derived from inbred line 478 (Fig. 4). Inbred line 478 was developed from the cross between inbred lines 8112 and 5003. Inbred lines 8112 and 5003 were selected from US commercial hybrids

3382 and 3147, respectively, in the 70s of the last century (Teng et al. 2004; Lai et al. 2010). It should be noted that inbred line 5003 has derived 24 elite F<sub>1</sub> hybrids and at least 30 elite inbred lines. Inbred line 8112 has derived at least 53 elite hybrids and at least 20 elite inbred lines (Li and Wang 2010). Therefore, commercial hybrids are the important sources for selecting elite inbred lines in developing countries.

In practice, the prediction can be based on the phenotypic data routinely generated by breeding programs, but the precision of prediction may be low. Genotype information of an F<sub>1</sub> can be deduced from its two inbred parents. To include genotypic information in prediction, inbred parental lines need to be genotyped first. In this study, inbred line per se performance was also used to determine the additive and dominance genetic effects of and QTL and markers. We expect the use of inbred line per se data will result in better estimations of these effects, which were then used in predictions.

Based on results in this study, we recommend that the genetic architecture of the trait of interest should be evaluated first, including analysis of genetic variance, estimate of heritability, combining ability analysis, QTL mapping, etc. Then the appropriate prediction model can be determined by the genetic architecture. If BV+GWP is chosen as the suitable model, genetic relationship between individuals as covariance matrix in mixed model need to be calculated by pedigree or molecular markers. After the genotypic values of untested combinations are predicted, the next step is to select the superior hybrids for field testing. In general, a superior hybrid for one trait may not be superior for other traits. An index by considering different breeding objectives may be necessary to select the best hybrids with high grain yield and satisfactory agronomic traits as well.

**Acknowledgments** This research was supported by National Hi-Tech Research and Development Program of China (2012AA101104-1). We thank Professor Rex Bernardo, Department of Agronomy and Plant Genetics, University of Minnesota for his constructive comments and suggestions to previous versions of this paper.

## References

- Allard RW (1960) Principles of plant breeding. Wiley, New York 485 pp
- Bernardo R (1994) Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci* 34:20–25
- Bernardo R (1995) Genetic models for predicting maize single-cross performance in unbalanced yield trial data. *Crop Sci* 35:141–147
- Bernardo R (1996a) Best linear unbiased prediction of maize single-cross performance. *Crop Sci* 36:50–56
- Bernardo R (1996b) Best linear unbiased prediction of the performance of crosses between untested maize inbreds. *Crop Sci* 36:872–876
- Bernardo R (1998) A model for marker-assisted selection among single crosses with multiple genetic markers. *Theor Appl Genet* 97:473–478
- Bernardo R (1999) Marker-assisted best linear unbiased prediction of single-cross performance. *Crop Sci* 39:1277–1282
- Bernardo R, Yu J (2007) Prospects for genomewide selection for quantitative traits in maize. *Crop Sci* 47:1082–1090
- 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
- Cockerham CC (1967) Prediction of double crosses from single crosses. *Der Züchter* 37:160–169
- Comstock RE, Robinson HF (1948) The components of genetic variance in populations. *Biometrics* 4:254–266
- Comstock RE, Robinson HF (1952) Estimation of average dominance of genes. In: Gowen JW (ed) Heterosis. ISU Press, Ames, pp 494–516
- De Los Campos G, Naya H, Gianola D, Crossa J, Legarra A, Manfredi E, Weigel K, Cotes JM (2009) Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* 182:375–385
- Frisch M, Thiemann A, Fu J, Schrag TA, Scholten S, Melchinger AE (2010) Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. *Theor Appl Genet* 120:441–450
- Gardner CO, Eberhart SA (1966) Analysis and interpretation of the variety cross diallel and related populations. *Biometrics* 22: 439–452
- Griffing B (1956a) Concept of general and specific combining ability in relation to diallel crossing systems. *Aust J Biol Sci* 9:463–493
- Griffing B (1956b) A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity* 10:31–50
- Hallauer AR (1990) Methods used in developing maize inbreds. *Maydica* 35:1–16
- Hallauer AR, Carena MJ, Filho JBM (2010) Quantitative genetics in maize breeding. Iowa State University Press, Ames
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Eco Notes* 2:618–620
- Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–447
- Henderson CR (1984) Applications of linear models in animal breeding. University of Guelph, Guelph
- Hua J, Xing Y, Wu W, Xu C, Sun X, Yu S, Zhang Q (2003) Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proc Nat Acad Sci* 100:2574–2579
- Jenkins MT (1934) Methods of estimating the performance of double crosses in corn. *J Am Soc Agron* 26:199–204
- Jordan DR, Tao Y, Godwin ID, Henzell RG, Cooper M, McIntyre CL (2003) Prediction of hybrid performance in grain sorghum using RFLP markers. *Theor Appl Genet* 106:559–567
- Kearsey MJ, Jinks JL (1968) A general method of detecting additive, dominance and epistatic variation for metrical traits. I: theory. *Heredity* 23:403–409
- Lai J, Li R, Xu X, Jin W, Xu M, Zhao H, Xiang Z, Song W, Ying K, Zhang M, Jiao Y, Ni P, Zhang J et al (2010) Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat Genet* 42:1027–1030
- Li Y, Wang T (2010) Germplasm base of maize breeding in China and formation of key parents in maize breeding in China (in Chinese with English abstract). *J Maize Sci* 18(5):1–8
- Li H, Ye G, Wang J (2007) A modified algorithm for the improvement of composite interval mapping. *Genetics* 175:361–374
- Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor Appl Genet* 120:151–161

- Maenhout S, Baets BD, Haesaert G (2010) Prediction of maize single-cross hybrid performance: support vector machine regression versus best linear prediction. *Theor Appl Genet* 120:415–427
- Melchinger AE (1999) Genetic diversity and heterosis. In: Coors JG, Pandey S (eds) *The genetics and exploitation of heterosis in crops*. ASA-CSSA, Madison, pp 99–118
- Melchinger AE, Geiger HH, Seitz G, Schmidt GA (1987) Optimum prediction of three-way crosses from single crosses in forage maize (*Zea mays* L.). *Theor Appl Genet* 74:339–345
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829
- Quaas RL, Pollak EJ (1980) Modified equations for sire models with groups. *J Dairy Sci* 54:1868–1872
- 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
- Smith OS (1986) Covariance between line per se and testcross performance. *Crop Sci* 26:540–543
- Tang J, Yan J, Ma X, Teng W, Wu W, Dai J, Dhillon BS, Melchinger AE, Li J (2010) Dissection of the genetic basis of heterosis in an elite maize hybrid by QTL mapping in an immortalized F<sub>2</sub> population. *Theor Appl Genet* 120:333–340
- Teng W, Cao J, Chen Y, Liu X, Jing X, Zhang F, Li J (2004) Analysis of maize heterotic groups and patterns during past decade in China (in Chinese with English abstract). *Scientia Agricultura Sinica* 37(12):1804–1811
- 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
- Wang J, van Ginkel M, Podlich D, Ye G, Trethowan R, Pfeiffer W, DeLacy IH, Cooper M, Rajaram S (2003) Comparison of two breeding strategies by computer simulation. *Crop Sci* 43:1764–1773
- Wang J, van Ginkel M, Trethowan R, Ye G, Delacy I, Podlich D, Cooper M (2004) Simulating the effects of dominance and epistasis on selection response in the CIMMYT Wheat Breeding Program using QuCim. *Crop Sci* 44:2006–2018
- Yu Y, Wang R, Shi Y, Song Y, Wang T, Li Y (2007) Genetic diversity and structure of the core collection for maize lines in China. *Maydica* 52:181–194
- Zaidi PH, Srinivasan G, Sanchez C (2003) Relationship between line per se and cross performance under low nitrogen fertility in tropical maize (*Zea mays* L.). *Maydica* 48:221–231
- Zhang L, Li H, Li Z, Wang J (2008) Interactions between markers can be caused by the dominance effect of quantitative trait loci. *Genetics* 180:1177–1190