ORIGINAL ARTICLE

Small ad hoc versus large general training populations for genomewide selection in maize biparental crosses

Sofía P. Brandariz[1] · Rex Bernardo[1]

Abstract

Key message **For genomewide selection in each biparental population, it is better to use a smaller ad hoc training population than a single, large training population.**

Abstract In genomewide selection, different types of training populations can be used for a biparental population made from homozygous parents (*A* and *B*). Our objective was to determine whether the response to selection (R) and predictive ability (r_{MP}) in an *A/B* population are higher with a large training population that is used for all biparental crosses, or with a smaller ad hoc training population highly related to the *A/B* population. We studied 969 biparental maize (*Zea mays* L.) populations phenotyped at four to 12 environments. Parent–offspring marker imputation was done for 2911 single nucleotide polymorphism loci. For 27 *A/B* populations, training populations were constructed by pooling: (1) all prior populations with *A* as one parent (*A/**, where * is a related inbred) and with *B* as one parent (**/B*) [general combining ability (GCA) model]; (2) *A/** or **/B* crosses only; (3) all **/** crosses (same background model, SB); and (4) all **/**, *A/**, and **/B* crosses (SB + GCA model). The SB model training population was 450–6000% as large as the GCA model training populations, but the mean coefficient of coancestry between the training population and *A/B* population was lower for the SB model (0.44) than for the GCA model (0.71). The GCA model had the highest R and r_{MP} for all traits. For yield, R was 0.22 Mg ha^{-1} with the GCA model and 0.15 Mg ha^{-1} with the SB model. We concluded that it is best to use an ad hoc training population for each *A/B* population.

Introduction

In maize (*Zea mays* L.) breeding, genomewide selection (or genomic selection) is typically performed among the progeny within a biparental cross. Suppose a biparental population is formed from the cross between two maize inbreds (*A* and *B*) that belong to the Iowa Stiff Stalk Synthetic (BSSS) heterotic group. A training population for the *A/B* cross, as well as for all other BSSS biparental crosses, can be made by pooling all prior biparental crosses that belong to the same BSSS genetic background. On the other hand, because genomewide selection is most effective when the training population is representative of the *A/B* population (Schulz-Streeck et al. 2012; Riedelsheimer et al. 2013; Jacobson

et al. 2014), the response to selection (R) and predictive ability (r_{MP}) within the *A/B* population may be higher if the training population includes only those prior biparental populations that are most related to the *A/B* cross.

Suppose *A/** is a prior biparental cross with *A* as a parent, **/B* is a biparental cross with *B* as a parent, and **/** is a biparental cross with parents that are different from *A* and *B* but have the same genetic background as *A* and *B*. These prior biparental crosses lead to three types of training populations to predict the performance of progeny within the *A/B* biparental cross. In the general combining ability (GCA) model, the training population is formed by pooling all prior *A/** populations and all prior **/B* populations (Jacobson et al. 2014). In the same background (SB) model, the training population is formed by pooling all prior **/** populations. In the SB + GCA model, all prior **/**, *A/**, and **/B* biparental crosses are pooled as the training population for the A/B biparental cross. Among 30 maize A/B populations, the mean *R* for grain yield, moisture, and test weight was higher with the GCA model than with the SB model and SB + GCA model (Jacobson et al. 2014). The GCA model therefore provides an easy rule for choosing an ad

✉ Rex Bernardo
bernardo@umn.edu

[1] Department of Agronomy and Plant Genetics, University of Minnesota, 411 Borlaug Hall, 1991 Upper Buford Circle, Saint Paul, MN 55108, USA

hoc training population that is smaller but is highly related to each *A/B* population undergoing genomewide selection.

In the Jacobson et al. (2014) study, the number of crosses (N_X) in the training population was kept constant between the GCA model and the SB model, and the number of lines (N) was kept generally similar between the two models. This assumption was needed to make an equal resource comparison between the two models. However, this assumption did not reflect the reality that N_X and N are naturally higher in the SB model than in the GCA model because a breeding program has more */* crosses than *A/** and */*B* crosses. In practice, all prior */* crosses can be used in the SB model. As shown later in this study, the training population in the GCA model may consist of 4500 lines from all prior *A/** and */*B* biparental crosses, but the training population in the SB model may consist of 50,000 lines. Because r_{MP} increases as N increases (Daetwyler et al. 2008, 2010; Endelman et al. 2014; Lian et al. 2014), the larger training population may compensate for the lower level of relatedness between the training population and *A/B* population in the SB model.

Choosing only those */* crosses that meet a specified threshold of similarity with the *A/B* population may increase the relatedness between the SB training population and the *A/B* population. Furthermore, the GCA model assumed that both *A* and *B* had previously been used as parents of biparental crosses. While a new inbred obtained from an external source (Parra and Hallauer 1996; Phillips 2010) may immediately be used as one of the parents of an *A/B* cross, the new inbred would not have prior *A/** or */*B* data available. Having either *A/** crosses only or */*B* crosses only may decrease the effectiveness of the GCA model.

Our main objective was to determine whether R and r_{MP} in an *A/B* population are higher with a single, large training population that is used for all biparental crosses (SB model), or with smaller ad hoc training populations that have a high level of relatedness with a given *A/B* population (GCA model). Our specific goals were to determine whether the usefulness of the GCA model is diminished in comparison with having prior data on only *A* or *B*, or in comparison with a SB model that has a larger N_X and N as well as different levels of similarity between the */* populations and the *A/B* population.

Materials and methods

Phenotypic and marker data

The data and procedures have been previously described (Jacobson et al. 2014, 2015a, b; Lian et al. 2014; Brandariz and Bernardo 2018) but are also briefly described here for the readers' convenience. Monsanto provided us with testcross phenotypic and SNP marker data for 969 biparental

maize populations. The populations were evaluated for grain yield (Mg ha^{-1} at 155 g H$_2$O kg^{-1}), moisture (g H$_2$O kg^{-1}), and test weight (kg hL^{-1}) at four to 12 environments in the USA from 2000 to 2008 (Jacobson et al. 2014). A total of 27 F$_2$ populations were selected as the *A/B* populations according to criteria described by Jacobson et al. (2014, 2015a). The lines in the *A/B* and training populations had the same inbred as the tester. The *A* and *B* parents belonged to the same heterotic group, whereas the tester belonged to the opposite heterotic group. Among the 969 biparental crosses, 485 *A/B* populations were in one heterotic group, whereas 484 *A/B* populations were in an opposite heterotic group. For a given trait, F$_2$ populations with nonsignificant ($p > 0.05$) h^2 estimates were excluded from the analysis.

The parents of the populations were genotyped with 2911 SNP markers, whereas the progeny in each of the 27 F$_2$ populations was genotyped with 25–123 markers. The genotypes at each locus were coded as 1 if the line was homozygous for the SNP allele from parent *A*, -1 if the line was homozygous for the SNP allele from parent *B*, and 0 if the line was heterozygous. Marker loci that were monomorphic between the two parental inbreds or that had a minor allele frequency less than 0.10 were excluded within each population (Lian et al. 2014; Jacobson et al. 2015a). The SNP data for the progeny were then imputed from the parental SNP data, based on the conditional probability of a non-observed marker genotype given the two flanking marker genotypes (Jacobson et al. 2015a). A previous study with the same data sets showed that marker imputation improved the predictive ability and response to selection, and that around 500 SNP markers were sufficient for grain yield and 1000 SNP markers were sufficient for moisture and test weight (Jacobson et al. 2015a).

Training populations

The training populations were constructed as follows: (1) GCA model, wherein all *A/** populations and all */*B* populations were pooled in the training population for predicting the performance of progeny in the *A/B* population; (2) *A/** populations only; (3) */*B* populations only; (4) SB model, wherein the training population comprised all available */* populations within each heterotic group; (5) SB model with the same N_X and a similar N as the GCA model (SB$_{Equal}$); (6) SB model with a coefficient of similarity greater than 0.60 between the */* crosses and the *A/B* population (SB$_{0.60}$); (7) SB model with a coefficient of similarity greater than 0.70 between the */* crosses and the *A/B* population (SB$_{0.70}$); and (8) SB + GCA model, wherein all */*, *A/**, and */*B* populations were pooled. The coefficient of similarity between the parents of an *A/B* population and training population was calculated as the simple matching coefficient across the SNP

loci (Sokal and Michener 1958), as described by Jacobson et al. (2015b).

Coefficient of coancestry

For each *A/B* population, we estimated the coefficient of coancestry among *A*, *B*, and the parents denoted by * to which *A* and *B* were crossed. As shown in "Results" section, these coefficients of coancestry (f_{AB}, f_{A*}, and f_{*B}) determined the coefficient of coancestry between an individual in the training population and an individual in the *A/B* population. The marker-based coefficient of coancestry between any two individuals (*X* and *Y*) was estimated as $f_{XY} = (S_{XY} - \Theta)/(1 - \Theta)$, where S_{XY} was the marker similarity between *X* and *Y*, and Θ was the probability that two individuals share alleles that are alike in state but are not identical by descent (Lynch 1988; Bernardo 1993, 2010). Given that SNP loci are biallelic, we assumed Θ was equal to 0.50. This value of Θ further assumed allele frequencies of 0.50 among unrelated individuals. Nevertheless, Θ was expected to remain close to 0.50 as long as marker allele frequencies ranged from about 0.40–0.60. With the latter allele frequencies, Θ was $[1 - 2(0.60)(0.40)] = 0.52$ instead of 0.50.

Response to selection and predictive ability

Marker effects were obtained for all 2911 SNP loci (non-imputed and imputed) by ridge regression–best linear unbiased prediction (RR–BLUP) as implemented in the rrBLUP package (Endelman 2011) in R software (R Core Team 2017). We used RR–BLUP because previous studies have shown that more complex models did not substantially improve the prediction accuracy for yield (Lorenzana and Bernardo 2009; Heffner et al. 2011; Heslot et al. 2012; Riedelsheimer et al. 2012). Marker effects were estimated separately within each cross for each trait according to procedures described by Jacobson et al. (2014). The performance of all *N* individuals in the *A/B* population was predicted as $\mathbf{y} = \mu \mathbf{1} + \mathbf{Xm}$, where **y** was an $N \times 1$ vector of predicted performance; μ was the estimated overall mean; **1** was an $N \times 1$ vector with elements equal to 1; **X** was an $N \times N_M$ incidence matrix with elements of 1, −1, and 0; and **m** was an $N_M \times 1$ vector of RR–BLUP marker effects averaged across the biparental populations in the training population, e.g., *A/** and **/B* populations in the GCA model, and **/** populations in the SB model (Jacobson et al. 2014). The *R* and r_{MP} were calculated for each *A/B* population. For each trait, *R* was calculated as the phenotypic mean of the 10% of lines with the best predicted performance minus the overall mean of the *A/B* population. The r_{MP} was calculated as the correlation between the marker-predicted and observed values for the progeny in each *A/B* population. A *t* test was used to determine whether the *R* and r_{MP} values across the test populations were significantly different ($p < 0.05$) among the training population models for each trait.

Results

Training population size, genetic similarity, and coancestry

The number of crosses (N_X) and lines (*N*) varied among the different types of training populations. The ranking of the models in terms of size of the training population (largest to smallest) was as follows: SB + GCA, SB, $SB_{0.60}$, $SB_{0.70}$, GCA and SB_{Equal}, and *A/** and **/B* (Fig. 1). The SB + GCA model had a mean (range in parentheses) N_X of 347 (338–357) and a mean *N* of 54,466 (53,406–55,627). The SB model had a mean N_X of 320 (289–352) and a mean *N* of 49,941 (45,422–54,691, Fig. 1). The GCA model had a mean N_X of 27 (5–59) and a mean *N* of 4525 (894–10,171). In terms of *N*, the SB model training population was 450–6000% as large as the training population in the GCA model. The *A/** model training population was 60% of the size of the GCA model training population, and the **/B* model training population was 40% of the size of the GCA model training population.

The marker similarity between the training population and the *A/B* population ($S_{TP,A/B}$) varied among the different models for constructing the training population. The $S_{TP,A/B}$ was highest in the GCA model, with a mean $S_{TP,A/B}$ (range in parentheses) of 0.80 (0.73–0.86, Fig. 1). The *A/** and **/B* models had a mean $S_{TP,A/B}$ of 0.80 (0.73–0.85). The SB_{Equal} model had the lowest $S_{TP,A/B}$, with a mean of 0.72 (0.63–0.77), but the $S_{TP,A/B}$ values were close among the SB_{Equal}, SB, $SB_{0.60}$, and SB + GCA models.

The coefficient of coancestry between the training population and the *A/B* population ($f_{TP,A/B}$) likewise varied among the different models. First, the marker-based coefficients of coancestry had a mean (range in parentheses) of 0.47 (0.21–0.67) for f_{AB}, 0.65 (0.47–0.94) for f_{A*}, and 0.70 (0.45–0.92) for f_{*B} in the GCA model. Second, we found that the expected value of $f_{TP,A/B}$ was $(1 + f_{AB} + f_{A*} + f_{*B})/4$ in the GCA model. The estimated $f_{TP,A/B}$ in the GCA model had a mean of 0.71 (0.59–0.80). Third, the marker-based coefficients of coancestry had a mean (range in parentheses) of 0.44 (0.26–0.55) for f_{A*}, and 0.43 (0.26–0.52) for f_{*B} in the SB model. Fourth, we found that the expected value of $f_{TP,A/B}$ was $(f_{A*} + f_{*B})/2$ in the SB model. The estimated $f_{TP,A/B}$ in the SB model had a mean of 0.44 (0.28–0.51). Fifth, we found that the expected value of $f_{TP,A/B}$ was $(1 + f_{AB} + f_{A*} + f_{*B})/4$ in the *A/** and **/B* models. This expected value was the same as that for the GCA model, and the estimated $f_{TP,A/B}$ had a mean of 0.73 (0.57–0.81) in the *A/** model and 0.67 (0.57–0.78) in
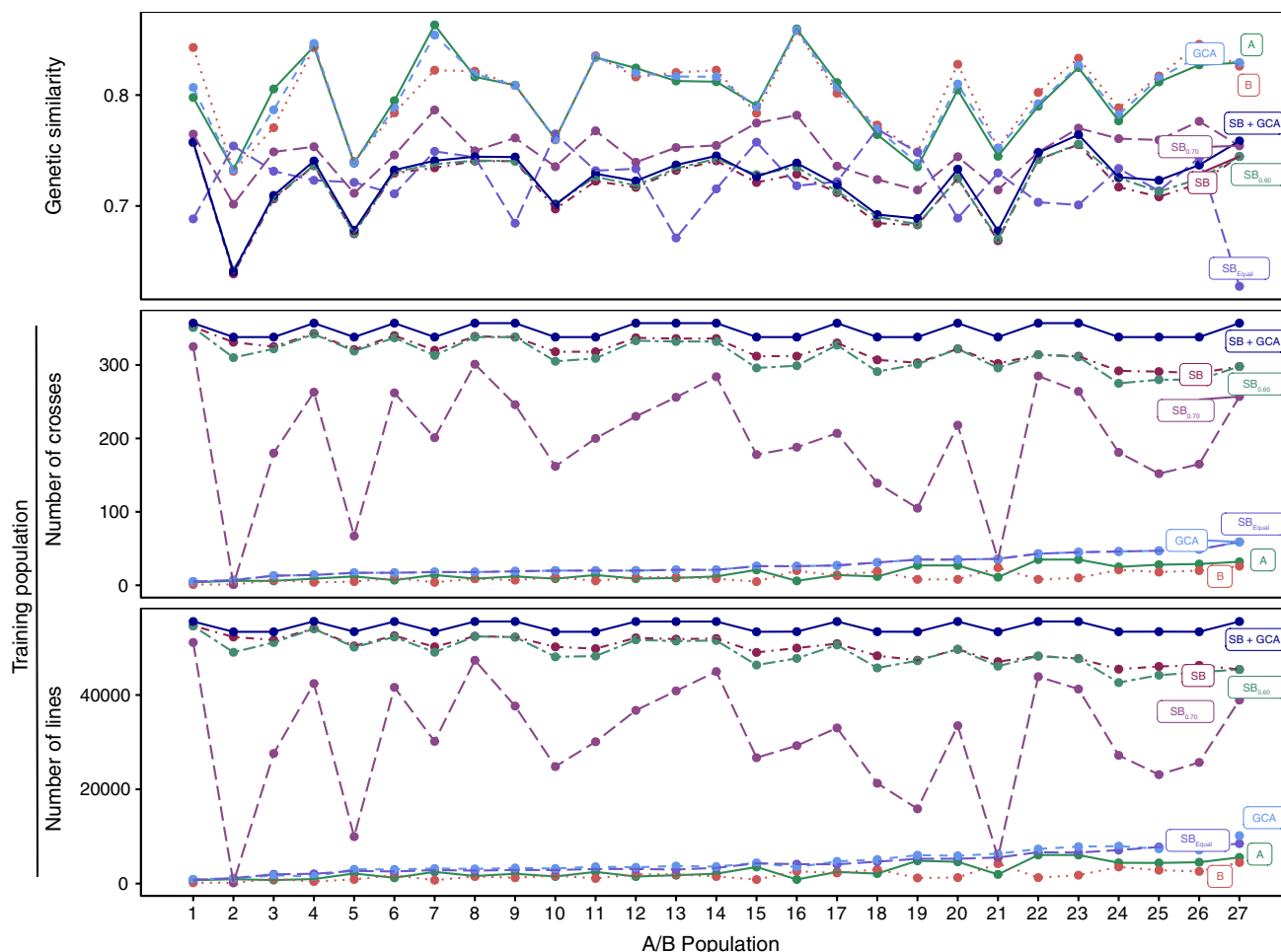
Fig. 1 Mean genetic similarity between training and *A/B* populations (top), number of crosses in the training population (middle), and number of lines in the training population (bottom) for eight training population models across the 27 *A/B* populations. Across the 27 *A/B* populations, the mean genetic similarity was: 0.80 for the GCA, *A/**, and **/B* models; 0.75 for the $SB_{0.70}$ model; and 0.72 for the SB_{Equal}, SB, $SB_{0.60}$, and SB + GCA models. The mean number of crosses in the training population was: 347 for the SB + GCA model; 320 for the SB model; 314 for the $SB_{0.60}$ model; 198 for the $SB_{0.70}$ model; 27 for the GCA and SB_{Equal} models; 16 for the *A/** model; and 11 for the **/B* model. The mean number of lines in the training population was: 54,466 for the SB + GCA model; 49,941 for the SB model; 48,992 for the $SB_{0.60}$ model; 30,774 for the $SB_{0.70}$ model; 4525 for the GCA model; 4175 SB_{Equal} model; 2726 for the *A/** model; and 1777 for the **/B* model

the **/B* model. Sixth, the estimated $f_{TP,A/B}$ in the SB + GCA model had a mean of 0.45 (0.28–0.53).

Response to selection and predictive ability with different training populations

A single, large training population (SB model and SB + GCA model) led to lower mean *R* and mean r_{MP} across the 27 *A/B* populations compared to the GCA model (Table 1). The differences in *R* and r_{MP} between the GCA model and SB model were statistically significant ($p < 0.05$) for moisture and test weight, whereas the differences between the GCA model and SB + GCA model were statistically significant only for test weight. For grain yield, *R* was 0.22 Mg ha^{-1} with the GCA model, 0.17 Mg ha^{-1} with the SB + GCA model, and 0.15 Mg ha^{-1} with the SB model. The corresponding r_{MP} for grain yield was 0.20 with the GCA model, 0.18 with the SB + GCA model, and 0.16 with the SB model. For moisture and test weight, the SB model and SB + GCA model led to larger decreases in *R* and r_{MP} (Table 1).

Restricting the SB model to having the same number of randomly selected crosses (SB_{Equal}) as the GCA model significantly decreased *R* and r_{MP} for all traits ($p < 0.05$, Table 1). For grain yield, *R* was initially 0.22 Mg ha^{-1} with the GCA model, decreased to 0.15 Mg ha^{-1} with the SB model, and decreased further to 0.11 Mg ha^{-1} with the SB_{Equal} model. The corresponding r_{MP} for grain yield was 0.20 with the GCA model, 0.16 with the SB model, and 0.12 with the SB_{Equal} model. Larger reductions were found for moisture and test weight (Table 1). The values of *R* and

Table 1 Mean and range (in parentheses) of response to selection (R) and predictive ability (r_{MP}) for grain yield, moisture, and test weight across 27 A/B maize populations, for the following training population models: (1) general combining ability (GCA) model with A/* and */B crosses; (2) A/* crosses only (A/*); (3) */B crosses only (*/B); (4) same background model with all */* crosses (SB); (5) SB with the same number of randomly selected */* crosses as the GCA model (SB$_{Equal}$); (6) SB with a coefficient of similarity between the */* crosses and A/B population greater than 0.60 (SB$_{0.60}$); (7) SB with a coefficient of similarity between the */* crosses and A/B population greater than 0.70 (SB$_{0.70}$); and (8) SB+GCA model with */*, A/*, and */B crosses

Model	Grain yield		Moisture		Test weight	
	R (Mg ha^{-1})	r_{MP}	R (g kg^{-1})	r_{MP}	R (kg hL^{-1})	r_{MP}
GCA	0.22a[a] (−0.16, 0.45)	0.20a (−0.04, 0.36)	−6.41a (−15.07, 0.33)	0.42a (−0.11, 0.67)	0.52a (0.10, 1.01)	0.36a (−0.06, 0.59)
A/*	0.15ab (−0.10, 0.36)	0.15ab (−0.04, 0.43)	−5.23ab (−12.01, 1.03)	0.35ab (−0.13, 0.63)	0.42ab (0.07, 0.73)	0.28b (−0.04, 0.47)
*/B	0.16ab (−0.15, 0.47)	0.16ab (−0.09, 0.39)	−5.17ab (−14.19, 0.44)	0.35ab (−0.02, 0.64)	0.34bcd (−0.41, 0.65)	0.26b (−0.17, 0.43)
SB	0.15ab (−0.26, 0.38)	0.16ab (−0.05, 0.42)	−4.38b (−9.36, −0.98)	0.31bc (−0.03, 0.58)	0.31cd (−0.05, 0.74)	0.24bc (−0.07, 0.43)
SB$_{Equal}$	0.11b (−0.20, 0.38)	0.12b (−0.10, 0.39)	−3.84b (−10.69, 2.36)	0.25c (−0.06, 0.46)	0.23d (−0.33, 0.53)	0.18c (−0.02, 0.38)
SB$_{0.60}$	0.15ab (−0.26, 0.38)	0.16ab (−0.05, 0.43)	−4.31b (−9.36, −0.98)	0.31bc (−0.03, 0.58)	0.31cd (−0.05, 0.74)	0.24bc (−0.06, 0.43)
SB$_{0.70}$	0.16ab (−0.19, 0.42)	0.16ab (−0.04, 0.41)	−4.40b (−10.33, −0.13)	0.32bc (−0.02, 0.58)	0.32bcd (−0.05, 0.73)	0.24bc (−0.06, 0.44)
SB+GCA	0.17ab (−0.30, 0.41)	0.18ab (−0.05, 0.45)	−5.00ab (−13.53, −0.98)	0.35ab (0.01, 0.60)	0.40bc (−0.05, 0.96)	0.29b (−0.07, 0.48)

[a]Within a column, estimates with a common letter were not significantly different ($p > 0.05$)

r_{MP} with the SB$_{Equal}$ model differed from those reported by Jacobson et al. (2014) because of three reasons: (1) the previous study used unimputed marker data, while in this study marker data were imputed from lower-density screening of the progeny in each biparental cross (see "Materials and methods"); (2) the previous study used 30 instead of 27 A/B populations; and (3) the crosses for the SB$_{Equal}$ training population were selected at random, which meant that the */* crosses used were not the same in the two studies.

Filtering the */* crosses (used in the SB model) to increase similarity with the A/B population (SB$_{0.60}$ or SB$_{0.70}$) was ineffective for improving R and r_{MP} compared to the SB model (Table 1). For grain yield, R was 0.22 Mg ha^{-1} with the GCA model and 0.15–0.16 Mg ha^{-1} with the SB, SB$_{0.60}$, and SB$_{0.70}$ models. Similar trends were found for moisture and test weight (Table 1).

When the training population included only the A/* crosses or only the */B crosses, R and r_{MP} were reduced although the differences were significant ($p < 0.05$) only for test weight (Table 1). For grain yield, R was 0.22 Mg ha^{-1} with the GCA model and 0.15–0.16 Mg ha^{-1} with the A/* and */B models, whereas the corresponding r_{MP} was 0.20 with the GCA model and 0.15–016 with the A/* and */B model. Similar trends were found for moisture and test weight (Table 1).

Discussion

Our main finding was that in genomewide selection in maize, it is better to use an ad hoc training population for each A/B biparental population (GCA model) than to use a single, large training population for all A/B populations (SB model and SB+GCA model). The results also showed the importance of both the relatedness between the training population and A/B population, and the size of the training population.

The SB model and SB+GCA models had, to our knowledge, the largest training populations ever described in the literature for plants (mean $N > 50,000$; Fig. 1). Despite this very large N in the SB model and SB+GCA model, the R and r_{MP} were higher with the GCA model, which had a mean N of about 4500. This result indicated that the relatedness between the A/B population and training population is more important than the size of the training population. Previous studies likewise highlighted the importance of including related crosses in the training population rather than increasing N by adding unrelated or less related crosses (Riedelsheimer et al. 2013; Jacobson et al. 2014; Lorenz and Smith 2015).

The higher relatedness in the GCA model was evidenced by the higher coefficient of coancestry between the training population and the A/B population ($f_{TP,A/B}$) in the GCA

model (0.71) than in the SB model (0.44) and SB + GCA model (0.45). In accordance with theoretical expectations, the estimated $f_{TP,A/B}$ was equal between the GCA model and the A/* or */B models. Furthermore, $f_{TP,A/B}$ is expected to be highest when individuals in an A/B population are used as the training population for other individuals in the same A/B population [i.e., A/B model, Jacobson et al. (2014)]. For the 27 A/B populations used in this study, the estimated $f_{TP,A/B}$ for the A/B model had a mean (range in parentheses) of 0.74 (0.60–0.84). The mean $f_{TP,A/B}$ of 0.71 in the GCA model was therefore close to the value of $f_{TP,A/B}$ in the A/B model.

The expected values of $f_{TP,A/B}$ were $(1 + f_{AB} + f_{A*} + f_{*B})/4$ in the GCA model and A/* and */B models, versus $(f_{A*} + f_{*B})/2$ in the SB model. The high $f_{TP,A/B}$ in the GCA and A/* and */B models was partly due to the higher values of f_{A*} and f_{*B} in these three models than in the SB model. These higher values of f_{A*} and f_{*B} in the GCA model and A/* and */B models were likely due to A/B crosses being made primarily within subgroups of parental inbreds. Suppose that the Iowa Stiff Stalk Synthetic maize heterotic group includes three subgroups: B14-type inbreds, B37-type inbreds, and B73-type inbreds. Inbreds within each subgroup are more closely related than inbreds in different subgroups. Furthermore, suppose that A/B crosses are most often (but not always) made within each subgroup. If A and B are both B73-type inbreds, the A/* and */B populations that are pooled into the GCA model training population will be mostly of the B73 background. In contrast, the SB model and SB + GCA model for a B73-type A/B population will also include crosses within the B14 and B37 subgroups. The f_{A*} and f_{*B} will consequently be lower in the SB model than in the GCA model, as was observed in this study.

The GCA model and the A/*, */B, and SB + GCA models obviously assume that the A and B parents have been used as parents of prior biparental crosses. There may be situations in which neither parent has been used in inbred development in previous years. In this situation, the GCA model cannot be used and the SB model will need to be used. The lower R and r_{MP} with the SB model than with the GCA model underscores the difficulty in predicting the performance of progeny of two untested parental inbreds.

Smaller sizes of the training population have been shown to decrease prediction accuracy (r_{MG}), which is defined as the correlation between predicted and true genotypic values and is equal to r_{MP}/h, where h is the square root of heritability (Daetwyler et al. 2008, 2010; Endelman et al. 2014; Lian et al. 2014). The expected prediction accuracy is $E(r_{MG}) = r^2 [(Nh^2/r^2 Nh^2 + M_e)]^{1/2}$, where r^2 is the linkage disequilibrium between a marker and a quantitative trait locus, and M_e is the effective number of chromosome segments (Lian et al. 2014). The above equation indicates that there is no minimum N required to reach a certain r_{MG}, because the required N will depend on both h^2 and M_e (Combs and Bernardo

2013). The equation for $E(r_{MG})$ applies only if the training population is genetically identical to the population undergoing genomewide selection (i.e., A/B model). For the 969 maize biparental crosses used in this study, Lian et al. (2014) reported mean values of $r^2 = 0.46$, $h^2 = 0.46$, and $M_e = 82$ for grain yield. The mean N in the SB model (49,941) was about 10 times larger than the mean N (4525) in the SB_{Equal} model. However, $E(r_{MG})$ [with the r^2, h^2, and M_e values reported by Lian et al. (2014)] increases by less than 4% when N increases from 4525 to 49,941. In contrast, if N increases tenfold from 450 to 4500, $E(r_{MG})$ increases by 31%. Therefore, when N is already large, the increase in r_{MG} due to further increases in N is minor. The training population was also already large with the A/* model (mean N of 2726) and */B model (mean N of 1777). This phenomenon of diminishing returns when N is large also explained why a 50% decrease in N led to less than a 50% decrease in r_{MP} when the training population included prior crosses for only one of the two parents (A/* and */B models) in comparison with the GCA model.

In the $SB_{0.60}$ and $SB_{0.70}$ models, filtering the */* crosses to increase the similarity between the training population and the A/B population ($S_{TP,A/B}$) increased $f_{TP,A/B}$ while maintaining a large N (mean $N > 36,000$, Fig. 1). However, R and r_{MP} were not significantly higher in the $SB_{0.60}$ and $SB_{0.70}$ models than in the SB model (Table 1). Because the mean $S_{TP,A/B}$ in the GCA model was 0.80, we tried a stricter threshold of 0.80 ($SB_{0.80}$ model) in 20 of the 27 A/B populations for which such a stricter threshold was possible. The $SB_{0.80}$ model had a mean (range in parentheses) N of 4983 (164, 12,483) and a mean N_X of 34 (1, 88), but the R and r_{MP} in the $SB_{0.80}$ model were lower than those in the $SB_{0.70}$ and $SB_{0.60}$ models (results not shown).

This lack of improvement was noteworthy because the mean N, N_X, and $S_{TP,A/B}$ were all roughly equal between the $SB_{0.80}$ model and the GCA model. Suppose A_1 and A_2 are the parents of A, and B_1 and B_2 are the parents of B. A grandparental training population can be formed by pooling the $A_1/*$, $A_2/*$, $*/B_1$, and $*/B_2$ biparental crosses, but such a grandparental training population was found ineffective for predicting the performance of progeny in the A/B biparental population (Hickey et al. 2014). While all of the individual alleles in the A/B population are found among the four grandparents, chromosomal blocks found in the A/B biparental population are represented better in the A/* and */B crosses than in the grandparental crosses.

These previous results (Hickey et al. 2014), as well as those in the current study, suggest that the usefulness of the GCA model may be due to having large blocks of chromosomes in common between the A/* and */B crosses used as the training population and the A/B population undergoing genomewide selection. The current study used data on 2911 SNP markers and the Hickey et al. (2014) study simulated

up to 100,000 markers. Given the agreement between our empirical results and the Hickey et al. (2014) simulation results, we speculate that our overall results would remain the same even if a larger number of SNP markers are used. In the maize populations we studied, the mean r^2 between adjacent SNP markers was 0.93 within each of the 27 *A/B* populations, 0.49 across the pool of *A/** and **/B* populations used in the GCA model, and 0.23 across the pool of **/** populations used in the SB model. Having a larger number of markers may increase the linkage disequilibrium in the training populations for both the GCA and SB models, but the effect of this higher linkage disequilibrium is unclear given that the linkage disequilibrium in the *A/B* populations was already very high (0.93) with 2911 SNP markers. The current study and the Hickey et al. (2014) studies both used RR–BLUP, and it remains to be seen whether the combined use of Bayesian models and higher marker densities in the SB model and SB + GCA model would improve R and r_{MP} for traits with no major QTL (such as grain yield) in maize biparental crosses.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bernardo R (1993) Estimation of coefficient of coancestry using molecular markers in maize. Theor Appl Genet 85:1055–1062

Bernardo R (2010) Breeding for quantitative traits in plants, 2nd edn. Stemma Press, Woodbury. ISBN 978-0-9720724-1-0

Brandariz SP, Bernardo R (2018) Maintaining the accuracy of genomewide predictions when selection has occurred in the training population. Crop Sci 58:1–6

Combs E, Bernardo R (2013) Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. Plant Genome. https://doi.org/10.3835/plantgenome2012.11.0030

Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3:e3395

Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA (2010) The impact of genetic architecture on genome-wide evaluation methods. Genetics 1031:1021–1031

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250–255

Endelman JB, Atlin GN, Beyene Y, Semagn K, Zhang X, Sorrells ME, Jannink J (2014) Optimal design of preliminary yield trials with genome-wide markers. Crop Sci 54:48–59

Heffner EL, Jannink JL, Sorrells ME (2011) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. Plant Genome 4:1–11

Heslot N, Yang HP, Sorrells ME, Jannink JL (2012) Genomic selection in plant breeding: a comparison of models. Crop Sci 52:146–160

Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R, Prasanna BM, Grondona M, Zambelli A, Windhausen VS, Mathews K, Gorjanc G (2014) Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Sci 54:1476–1488

Jacobson A, Lian L, Zhong S, Bernardo R (2014) General combining ability model for genomewide selection in a biparental cross. Crop Sci 54:895–905

Jacobson A, Lian L, Zhong S, Bernardo R (2015a) Marker imputation before genomewide selection in biparental maize populations. Plant Genome 8:1–9

Jacobson A, Lian L, Zhong S, Bernardo R (2015b) Minimal loss of genetic diversity after genomewide selection within biparental maize populations. Crop Sci 55:783–789

Lian L, Jacobson A, Zhong S, Bernardo R (2014) Genomewide prediction accuracy within 969 maize biparental populations. Crop Sci 54:1514–1522

Lorenz AJ, Smith KP (2015) Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. Crop Sci 55:2657–2667

Lorenzana RE, Bernardo R (2009) Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120:151–161

Lynch M (1988) Estimation of relatedness by DNA fingerprinting. Mol Biol Evol 5:584–599

Parra JR, Hallauer AR (1996) Utilization of exotic maize germplasm. In: Janick J (ed) Plant breeding reviews, vol 14. Wiley, Oxford, pp 165–187

Phillips RL (2010) Mobilizing science to break yield barriers. Crop Sci 50:S-99–S-108

Riedelsheimer C, Technow F, Melchinger AE (2012) Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. BMC Genom 13:452

Riedelsheimer C, Endelman JB, Stange M, Sorrells ME, Jannink J-L, Melchinger AE (2013) Genomic predictability of interconnected biparental maize populations. Genetics 194:493–503

Schulz-Streeck T, Ogutu JO, Karaman Z, Knaak C, Piepho HP (2012) Genomic selection using multiple populations. Crop Sci 52:2453–2461

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 38:1409–1438

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/