Genomewide Selection and Marker-Assisted Recurrent Selection in Doubled Haploid versus F_2 Populations

Patricio J. Mayor and Rex Bernardo*

ABSTRACT

Molecular markers have been previously found useful for increasing genetic gain in maize (*Zea mays* L.). The use of doubled haploids (DHs) instead of F_2 plants (or, equivalently, F_3 families) may permit a better estimation of marker–trait associations. Our objective was to determine the usefulness of DH versus F_2 populations in marker-assisted recurrent selection (MARS) and genomewide selection. We simulated testcrosses from a DH population and an F_2 population from the same cross between two inbreds and studied genetic models defined by the number of quantitative trait loci (QTL) and trait heritability (*H*). Equal-time comparisons of selection response were between Cycle 3 with an F_2 population and Cycle 2 with a DH population. For the genetic model of 100 QTL, $H = 0.20$ and a population size of $N = 100$, the ratio of response to selection in DH versus F_2 populations was $R_{DH:F2} = 109\%$ for genomewide selection and 128% for MARS. For the genetic model of 20 QTL, $H = 0.80$ and $N = 100$ these values decreased to 99% for genomewide selection and 109% for MARS. Although genomewide selection was superior to MARS for a given type of population, the advantage of using DH instead of F_2 populations was greater in MARS than in genomewide selection. We concluded that DH populations are most useful in genomewide selection and MARS when many QTL control the trait, *H* is low, and *N* is small.

Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, 411 Borlaug Hall, 1991 Upper Buford Cir., St. Paul, MN 55108. Received 6 Oct. 2008. *Corresponding author (bernardo@umn.edu).

Abbreviations: DH, doubled haploid; MARS, marker-assisted recurrent selection; QTL, quantitative trait loci; $R_{DH:F2}$, ratio between the cumulative response to selection in the DH population and the cumulative response to selection in the F_2 population; $R_{GWS:MARS}$, ratio between the cumulative response to genomewide selection and the cumulative response to MARS; V_G, genetic variance; V_M, variance explained by markers; V_R, residual variance.

More than a half-century after the use of doubled haploids (DHs) in maize (*Zea mays* L.) breeding was first proposed (Chase, 1951), DH technology is now routinely used in maize inbred development (Seitz, 2005). Producing DH lines requires four steps: (i) inducing haploids by crossing heterozygous plants as maternal parents with a DH inducer; (ii) identifying haploid kernels through morphological markers; (iii) doubling chromosomes of putative haploids by colchicine treatment; and (iv) selfing to obtain seeds of DH lines (Chase, 1951; Bordes et al., 1997; Melchinger et al., 2005; Seitz, 2005). The DH technology enhances maize breeding in two ways. First, it reduces the time required to produce inbreds. Whereas six or more generations of selfing are needed to produce inbreds from an F_1, DH technology produces inbreds in only two generations. Second, the effectiveness of selection is improved by the higher genetic variance among DH lines than among F_2 plants or selfed (e.g., F_3 or F_4) families (Gallais, 1990; Seitz, 2005; Gallais and Bordes, 2007).

In addition to DH technology, molecular markers are now routinely used in selection by commercial maize breeders (Johnson,

Published in Crop Sci. 49:1719–1725 (2009).
doi: 10.2135/cropsci2008.10.0587
© Crop Science Society of America
677 S. Segoe Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

2004; Eathington et al., 2007). Two approaches for marker-based selection have been applied in maize. First, molecular markers have been used to map major quantitative trait loci (QTL) and subsequently introgress these major QTL into other inbreds (Concibido et al., 2004; Pumphrey et al., 2007). This approach relies on finding QTL with large effects that are consistent across different genetic backgrounds. Second, several cycles of marker-based selection have been used to improve the mean performance of F_2 populations before inbreeding (Edwards and Johnson, 1994; Johnson, 2004). In this procedure, called marker-assisted recurrent selection (MARS), marker–trait associations are first identified from field trials of testcrosses of F_2 plants or F_3 families (Cycle 0). The best Cycle 0 families are selected based on an index that combines testcross phenotypic data and marker scores (Lande and Thompson, 1990). Next, two to three cycles of selection are performed based solely on markers with significant effects. Such marker-based selection is conducted in an off-season nursery (e.g., in Hawaii or Puerto Rico) or in a greenhouse, where multiple generations can be grown in a year and but where phenotypic evaluations are not meaningful (Johnson, 2004). Empirical results have shown that MARS is effective for multiple-trait improvement of maize populations before inbreeding (Johnson, 2004; Eathington et al., 2007). In contrast to MARS, which utilizes only those markers with significant effects, genomewide selection involves calculating breeding values for all markers and utilizes all markers in selection without testing them for their significance (Meuwissen et al., 2001). Simulation studies have indicated that the response to genomewide selection is 18 to 43% higher than the response to MARS (Bernardo and Yu, 2007).

The use of maize DH populations instead of F_2 populations (i.e., F_2 plants or F_3 families) may permit a better estimation of effects associated with markers (Hyne et al., 1995) in MARS and genomewide selection. While Bernardo and Yu (2007) studied the response to genomewide selection with DH populations, studies have not been published on the usefulness of DH versus F_2 populations in MARS and in genomewide selection. Our objective in this research was to determine, by computer simulation, the usefulness of DH versus F_2 populations in MARS and genomewide selection. We examined the influence of number of QTL, trait heritability, and population size on the utility of DH versus F_2 populations.

MATERIALS AND METHODS

Populations and Genetic Models

Two parental inbreds were crossed in each repeat of the simulation experiments. The F_1 plants were crossed to a haploid inducer to generate maternal haploids. Chromosomes were then doubled to produce DH lines. Simultaneously, F_1 plants were selfed to form F_2 plants. A total of N DH lines and N F_2 plants were crossed to an unrelated inbred tester.

Each simulation experiment comprised a combination of a population size (N = 100, 250, or 450) and three different genetic models defined by the number of segregating QTL (L) and trait heritability (H) on a testcross-mean basis across environments. The three genetic models were 100 QTL, H = 0.20; 60 QTL, H = 0.50; and 20 QTL, H = 0.80. Each repeat of a simulation experiment differed in the location of QTL, genotypes of sampled individuals, and phenotypic values. We simulated 1000 repeats of each experiment.

The linkage map of 10 maize chromosomes comprised 1749 cM (Senior et al., 1996). A total of 100 equally spaced markers were used in MARS in accordance with preliminary studies we conducted (data not shown) and with previous studies indicating that given the values of N we considered, the response to MARS is largest when 64 to 128 markers were used (Bernardo and Yu, 2007). A total of 200 equally spaced markers were used in genomewide selection in accordance with preliminary studies we conducted (data not shown) and with previous studies (Bernardo and Yu, 2007).

Testcross genetic effects of individual QTL were modeled according to a geometric series (Lande and Thompson, 1990; Bernardo and Yu, 2007) so that few QTL had large effects and many QTL had smaller effects. At the kth QTL, the testcross effect of the favorable allele was a^k, with $a = (L - 1)/(L + 1)$ (Lande and Thompson, 1990). The testcross effect of the less favorable allele was $-a^k$. Therefore, the first QTL had the largest effect, the second QTL had the second largest effect, and the Lth QTL had the smallest effect. The first parental inbred had the favorable alleles at the even-numbered QTL and the less-favorable allele at the odd-numbered QTL. Given that the QTL were located randomly in the genome in each repeat of a simulation experiment, coupling and repulsion linkages among QTL were therefore generated at random without regard for the magnitude of the QTL effects. Because testcross means behave in an additive manner regardless of the level of dominance (Hallauer and Miranda, 1981), dominance was assumed absent. Epistasis was also assumed absent.

Random nongenetic effects were added to the genotypic values to obtain testcross phenotypic values. These random effects had a normal distribution with mean of zero and were scaled so that broad-sense heritability among testcrosses was H = 0.20, 0.50, or 0.80 in the initial F_2 populations. In this article, H refers to the testcross-mean heritability in the F_2 population rather than the DH population. Given that the testcross genetic variance (V_G) in a DH population is twice as large as the V_G in the F_2 population, H = 0.20, 0.50, or 0.80 in the F_2 population corresponded to H = 0.33, 0.67, and 0.89 in the DH population.

Simulation of MARS

The MARS procedure with F_2 populations involved one cycle of selection based on both phenotypic data and marker scores followed by three cycles of selection based on marker scores only. In Cycle 0, testcrosses of N F_2 individuals were evaluated in six environments. Replication across environments allowed the estimation of V_G and residual variance (V_R) in genomewide selection by best linear unbiased prediction (Meuwissen et al., 2001). The variance components were estimated by equating the observed mean squares in an analysis of variance to their expectations and solving for the desired variance component.

Markers with significant effects were detected only in Cycle 0 (i.e., when phenotypic data were available) as described by Bernardo and Charcosset (2006). First, multiple regression of phenotypic values on the number of marker alleles from the first parental inbred was performed for each chromosome. Backward elimination was

used to identify the significant markers on each chromosome. The significance levels used were $\alpha = 0.20$, 0.30, or 0.40, according to previous studies (Hospital et al., 1997; Moreau et al., 1998) indicating that the response to MARS is greater with relaxed significance levels. The results presented in this manuscript are for the significance level for which the response to MARS was largest. Next, all markers found significant in the per-chromosome analysis were fitted in a multiple-marker model. Marker scores for each of the N F_2 plants were calculated as $M_j = \sum b_i X_{ij}$, where M_j was the marker score of the jth individual; b_i was the weight for the ith marker locus; and X_{ij} was equal to 1 if the jth individual was homozygous for the marker allele (at the ith marker locus) from the first parental inbred, −1 if the individual was homozygous for the marker allele from the second parental inbred, and 0 if the individual was heterozygous. The b_i weights were obtained from multiple regression of trait values on X_{ij} (Lande and Thompson, 1990; Hospital et al., 1997). Marker scores were then combined with the testcross performance of each F_2 plant in a least-squares index (Lande and Thompson, 1990) with constraints added so that the phenotypic score always had a positive contribution to the index (Hospital et al., 1997).

The $N_{Sel} = 6$ individuals with the highest least-squares index values (Lande and Thompson, 1990) in Cycle 0 were random-mated to form Cycle 1. Selection in Cycle 1 was then based on marker scores only and the best $N_{Sel} = 6$ plants were random-mated to form Cycle 2. Procedures in Cycle 1 were then repeated in Cycles 2 and 3.

The MARS procedure with DH populations was the same as the MARS procedure with F_2 populations except for (i) the type of progenies used in Cycle 0 (DH) and (ii) having one less cycle of marker-based selection with DH populations than with F_2 population. Crossing selected F_2 plants in Cycle 0 immediately leads to a segregating population (Cycle 1) that can be subjected to selection. In contrast, crossing selected DH lines in Cycle 0 leads to a mixture of nonsegregating F_1s that needed to be recombined a second time to form a segregating Cycle 1 population, thereby adding one season to the selection process (Fig. 1). An equal-time comparison was therefore between F_2 Cycle 3 and DH Cycle 2.

Responses to selection were expressed in units of the genetic standard deviation among testcrosses of F_2 plants in Cycle 0. The variance of the response was determined from the 1000 repeats of the simulation experiments. Differences in selection response were tested with z-tests ($P = 0.05$).

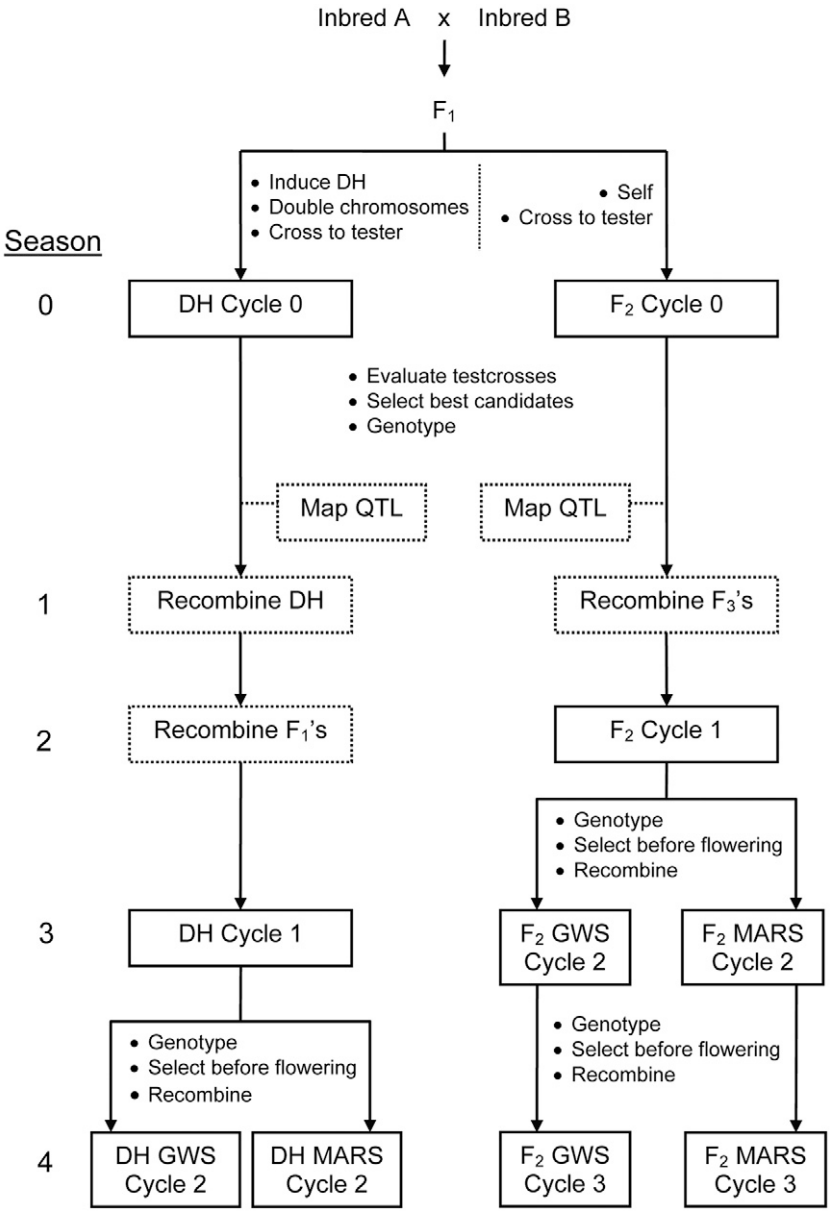

Figure 1. Breeding schemes for genomewide selection (GWS) and marker-assisted recurrent selection (MARS) with doubled haploid (DH) and F_2 testcross populations.

Simulation of Genomewide Selection

Procedures for genomewide selection were identical to procedures for MARS except for the following two ways. First, selection was based only on testcross phenotypic performance in Cycle 0 of genomewide selection. This procedure was used because previous results (Bernardo and Yu, 2007) showed that combined phenotypic and marker selection in Cycle 0 did not improve the response. Second, all markers were used in calculating marker scores in genomewide selection. Specifically, best linear unbiased predictions of breeding values associated with each marker were obtained in genomewide selection. The performance of the Cycle 0 testcrosses was modeled as $\mathbf{y} = \mu\mathbf{1} + \mathbf{Xg} + \mathbf{e}$, where $\mathbf{y}$ was an $N \times 1$ vector of phenotypic testcrosses means in Cycle 0; μ was the overall mean; $\mathbf{1}$ was an $N \times 1$ incidence vector with all elements equal to 1; $\mathbf{X}$ was an $N \times 200$ design matrix with elements equal to 1 if the individual in Cycle 0 was homozygous for the marker allele from the first parental inbred, −1 if it was homozygous for the marker allele from the second parental inbred, and 0 if it was heterozygous; $\mathbf{g}$ was a 200×1 vector of breeding values associated with the marker allele from the first parental inbred; and $\mathbf{e}$ was an $N \times 1$ vector of residual effects.

The variance of the breeding values for each of the $L = 200$ marker loci was assumed equal to the estimated V_G divided by L (Meuwissen et al., 2001). The $\mathbf{g}$ vector was obtained by solving the mixed-model equations (Meuwissen et al., 2001) and was then used to calculate marker scores (Bernardo and Yu, 2007).

RESULTS AND DISCUSSION

Response to Selection in DH versus F_2 Populations: Genetic Models and Population Sizes

In terms of selection response, the advantage of using DH populations instead of F_2 populations varied with the number of QTL, heritability (*H*), selection procedure (genomewide selection vs. MARS), and population size (*N*). Overall, DH populations were most advantageous when many QTL controlled the trait, *H* was low, and *N* was small. For the 100 QTL, $H = 0.20$ genetic model with $N = 100$, the ratio between the cumulative response to selection in the DH population (Cycle 2) and the cumulative response to selection in the F_2 population (Cycle 3) was $R_{DH:F2} = 109\%$ for genomewide selection and 128% for MARS (Table 1). These $R_{DH:F2}$ values decreased to 99% with genomewide selection and 109% with MARS under the simpler genetic model of 20 QTL, $H = 0.80$ ($N = 100$). In this case, the difference in the response to genomewide selection in DH and F_2 populations was not significant ($LSD_{0.05} = 0.09$).

A large number of QTL and a low *H* lead to a low power for detecting QTL and poor precision in the estimates of QTL effects (Beavis, 1994). Given that DH lines increase the power and precision in QTL mapping (Hyne et al., 1995), $R_{DH:F2}$ was expected to be largest in the 100 QTL, $H = 0.20$ genetic model. In contrast, $R_{DH:F2}$ was expected to be lowest in the 20 QTL, $H = 0.80$ genetic model for which the power and precision in QTL mapping may be high to start with in an F_2 population. Furthermore, the accuracy of genomewide predictions of genetic merit depends on *H* (Calus et al., 2008; Daetwyler et al., 2008). If *H* is 0.20 in an F_2 population, the use of DH lines increases *H* to 0.33; that is, 65% increase in *H*. If *H* is 0.80 in an F_2 population, the use of DH lines increases *H* to 0.89; that is, 11% increase in *H*. The proportion of increase in *H* and, consequently, in the accuracy of genomewide predictions of genetic merit is therefore largest when *H* is low. This result explains the greater advantage of DH populations in genomewide selection when *H* is low than when *H* is high.

The increase in response due to the use of DH populations instead of F_2 populations was largest in the first cycle of selection (i.e., Cycle 0→1). This result was expected because the genetic variance is twice as large in DH populations than in F_2 populations (Gallais, 1990; Seitz, 2005). For the 100 QTL, $H = 0.20$ genetic model with $N = 100$, the increase in the Cycle 0→1 response when DHs instead of F_2 populations were used was 69% for genomewide selection and 77% for MARS (Fig. 2). For the 20 QTL, $H = 0.80$ genetic model with $N = 100$, the increase in the Cycle 0→1 response when DHs instead of F_2 populations were used was 45% for genomewide selection and 50% for MARS (Fig. 2b). At the end of genomewide selection or MARS, the additional cycle of selection among F_2 populations compared with DH populations (Fig. 1) did not compensate for the higher Cycle 0→1 responses with DHs than with F_2 populations (Fig. 2).

The advantage of selection in DHs instead of F_2 populations decreased as *N* increased (Table 1). For the 100 QTL, $H = 0.20$ genetic model, $R_{DH:F2}$ for MARS decreased from 128% with $N = 100$, to 118% with $N = 250$ and 111% with $N = 450$ (Table 1). While the response to genomewide selection for the 100 QTL, $H = 0.20$ genetic model with $N = 100$ was greater with DHs than with F_2 populations ($R_{DH:F2} = 109\%$), the corresponding differences in response were not significant when *N* increased to 250 or 450 (Table 1). For genomewide selection under both the 60 QTL, $H = 0.50$ and the 20 QTL, $H = 0.80$ genetic models, the advantage of using DH instead of F_2 populations disappeared when *N* increased from 100 to 250 or 450 (Table 1). For the 20 QTL, $H = 0.80$ genetic model, DH populations were likewise not advantageous for either genomewide selection or MARS when *N* increased from 100 to 250 or 450 (Table 1).

Lande and Thompson (1990) demonstrated that the product of *N* and *H* has an inverse relationship with the proportion of V_G explained by significant markers. In theory, $R_{DH:F2}$ values will decrease as the proportion of V_G explained by significant markers approaches 1.0. It was therefore expected that increases in either or both *N* and *H* would lead to a lower $R_{DH:F2}$ ratio. Furthermore, the response to MARS is proportional to *H* and the proportion of V_G explained by markers

Table 1. Response to genomewide selection and marker-assisted recurrent selection (MARS) in simulated doubled haploid (DH) and F_2 maize testcross populations of different sizes (*N*).

Genetic model	Method†	*N* = 100			*N* = 250			*N* = 450		
		DH	F_2	$R_{DH:F2}$‡	DH	F_2	$R_{DH:F2}$	DH	F_2	$R_{DH:F2}$
100 QTL, $H = 0.20$	GWS	2.37§	2.17	109%	3.04	2.97	102%	3.44	3.48	99%
	MARS	2.10¶	1.64	128%	2.77	2.34	118%	3.22	2.91	111%
	$R_{GWS:MARS}$	113%	132%		110%	127%		107%	120%	
60 QTL, $H = 0.50$	GWS	3.08	2.94	105%	3.79	3.78	100%	4.16	4.23	98%
	MARS	2.83	2.38	119%	3.55	3.34	106%	3.98	3.94	101%
	$R_{GWS:MARS}$	109%	124%		107%	113%		105%	107%	
20 QTL, $H = 0.80$	GWS	3.28	3.30	99%	3.84	3.91	98%	4.17	4.23	99%
	MARS	3.20	2.93	109%	3.84	3.72	103%	4.09	4.04	101%
	$R_{GWS:MARS}$	103%	113%		100%	105%		102%	105%	

†GWS, genomewide selection; $R_{GWS:MARS}$, ratio between the response to genomewide selection and the response to MARS.

‡$R_{DH:F2}$ is the ratio between the response in an F_2 population and the response in the DH population.

§Responses are in units of the testcross genetic standard deviation in Cycle 0 ($LSD_{0.05} \approx 0.09$). Responses are based on the mean of Cycle 2 for DH populations and the mean of Cycle 3 for F_2 populations (Fig. 1).

¶For MARS, responses are given for the significance level ($\alpha = 0.10, 0.20, 0.40$) that maximized the response.

(Lande and Thompson, 1990). However, the proportion of V_G explained by markers is difficult to assess in simulation studies because the estimated amount of the variance explained by markers (V_M) can exceed V_G. In our study, the values of V_M/V_G ranged from 2.11 for the genetic model of 100 QTL, $H = 0.20$ in an F_2 population with $N = 100$, to 1.01 for the same genetic model in a DH population with $N = 450$. This overestimation of V_M was due to a finite population size and the resulting upward bias in the magnitude of the estimated QTL effects (Beavis, 1994). The gross overestimation of V_M made it difficult to directly interpret the results in terms of the true proportion of V_G explained by markers.

Overall, the results for different genetic models and N were consistent with DH populations being most useful when marker effects are difficult to estimate. The literature has shown that obtaining precise estimates of QTL effects is difficult when the number of QTL is large, H is low, and N is small (Beavis, 1994; Melchinger et al., 1998), and these were the situations in which DH populations were most advantageous over F_2 populations in terms of the response to genomewide selection or MARS.

Response to Selection in DH versus F_2 Populations: Genomewide Selection versus MARS

Across genetic models and population sizes, the ratio between the cumulative response to genomewide selection and the cumulative response to MARS ($R_{GWS:MARS}$) ranged from 100 to 113% for DH populations (Table 1). The $R_{GWS:MARS}$ ratios ranged from 105 to 132% for F_2 populations. Regardless of the population size and the type of population used, the response to genomewide selection was larger ($LSD_{0.05} \approx 0.09$) than the response to MARS for the 100 QTL, $H = 0.20$ and the 60 QTL, $H = 0.50$ genetic models (Table 1). For the 20 QTL, $H = 0.80$ genetic model, differences in the response to genomewide selection and MARS in DH populations were nonsignificant regardless of N. For this genetic model, however, genomewide selection in F_2 populations remained superior to MARS in F_2 populations regardless of N.

While the responses to genomewide selection were greater than or equal to the corresponding responses to MARS, the advantage of using DH instead of F_2 populations was greater in MARS than in genomewide selection. Specifically, $R_{DH:F2}$ ranged from 98 to 109% with genomewide selection and from 98 to 128% with MARS. This result was consistent with DH populations being most useful when having precise estimates of marker effects is most crucial. In MARS, in which markers with significant effects ($\alpha = 0.20–0.40$) are used in selection, having good estimates of QTL effects is important because genetic gain is expected only from a subset of the QTL (Bernardo and Charcosset, 2006). But in genomewide selection, which models 100% of the V_G, having good estimates of QTL effects is arguably less crucial because genetic gain is spread out across a larger number of QTL (and conceivably, all QTL) affecting the trait.

Application in Breeding Programs

The results suggested that DH populations would be most useful for genomewide selection or MARS for traits such as grain yield and agronomic traits, which are controlled by many QTL and which often have a low H. In practice, improvement of several complex traits is sought in MARS (Edwards and Johnson, 1994; Johnson, 2001; Johnson, 2004; Eathington et al., 2007) and in genomewide selection in maize. Of the three genetic models we studied, the genetic model for which DH populations were most advantageous (100 QTL, $H = 0.20$) was arguably the most representative of the actual genetic models encountered in

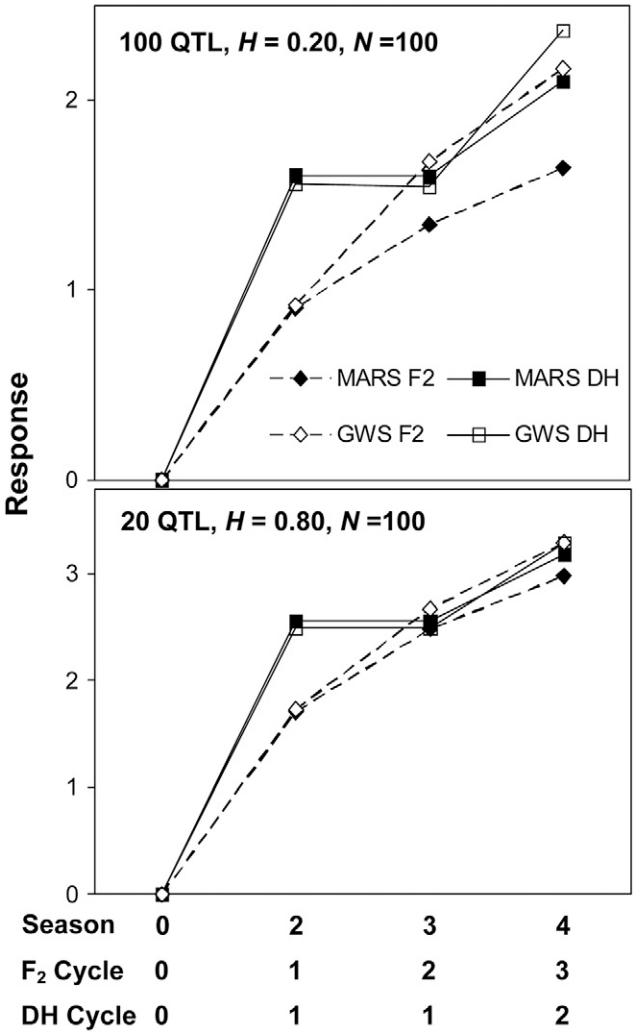

Figure 2. Response to genomewide selection (GWS) and marker-assisted recurrent selection (MARS) in doubled haploid (DH) and F_2 populations of different sizes (N) under genetic models with different numbers of QTL and trait heritability (H). Responses are in units of the testcross genetic standard deviation in Cycle 0. Season 1 is used for recombining selected Cycle 0 progenies in both F_2 and DH populations, and a second generation of recombination is required in DH populations (Fig. 1).

genomewide selection and MARS using a multiple-trait index in maize.

In maize, MARS with selfed families typically involves population sizes of N = 100 to 150 (Johnson, 2001). A practical consideration is whether current DH technology allows the routine production of N = 100 to 150 DH lines. Fifty years ago, Coe (1959) described "a line of maize with high haploid frequency" of 2.3%. The use of improved haploid inducers such as RWS (Röber et al., 2005), however, leads to a mean haploid induction rate of about 8%. The efficiency of chromosome doubling by colchicine treatment is approximately 10% (Melchinger et al., 2005). In a comparison of testcross selection between DH and S_2 populations, Seitz (2005) concluded that DH lines were advantageous with population sizes of N = 122 and 204. The current DH technology can already achieve population sizes of $N > 100$ (P. Mayor, unpublished data, 2008). The efficiency of DH induction will further increase as maize researchers continue to develop better haploid inducers and better techniques for chromosome doubling, for example, embryo rescue (G. Graham, personal communication, 2008).

The advantage of marker-based selection in genomewide selection or MARS is mainly in gain per unit time rather than in gain per cycle (Edwards and Johnson, 1994; Hospital et al., 1997; Bernardo and Yu, 2007). While DH lines are useful in Cycle 0 of genomewide selection or MARS, the development of DH lines in each cycle of genomewide selection or MARS will substantially increase the time needed for each cycle of selection. In the scheme we considered, three cycles of MARS with F_2 populations requires four seasons (Fig. 1). In contrast, developing DH lines after each of the three cycles of MARS would require seven seasons. However, DH lines may be developed after each cycle of genomewide selection or MARS to directly develop new inbreds for testcrossing and evaluation (rather than for forming the next cycle of selection).

The 13% maximum advantage in response to genomewide selection over MARS with DH populations in this study was less than the maximum advantage of 43% found by Bernardo and Yu (2007). This discrepancy is explained by a difference in the number of cycles of selection used. While Bernardo and Yu (2007) calculated the cumulative response to genomewide selection with DH populations based on the mean of Cycle 3, we calculated the response based on the mean of Cycle 2. Up to three cycles of selection are usually conducted in large-scale applications of MARS in maize (Eathington et al., 2007). Given this usual practice and given that the use of DH instead of F_2 populations in Cycle 0 leads to one less cycle of selection, we therefore considered selection only until Cycle 2 for DH populations. Including a third cycle of selection for DH populations will delay the testcross evaluations by 1 yr (Fig. 1). Having three cycles of selection with DH populations may be worthwhile only if Cycle 3 will be immediately evaluated for per se performance in a nursery (as assumed by Bernardo and Yu [2007]) instead of for testcross performance in field trials (as we assumed in this study), or if the cycle time can be sped up through early harvesting (e.g., 35 d after pollination), through embryo rescue, or with early maturing maize germplasm.

In conclusion, our results indicated that DH populations are useful in genomewide selection and MARS for traits controlled by many QTL and with a low heritability. We are currently conducting field studies in maize to validate these simulation results.

References

Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250–266. *In* Proc. 49th Annu. Corn Sorghum Ind. Res. Conf. Chicago, IL. 7–8 Dec. 1994. Am. Seed Trade Assoc., Washington, DC.

Bernardo, R., and A. Charcosset. 2006. Usefulness of gene information in marker-assisted recurrent selection: A simulation appraisal. Crop Sci. 46:614–621.

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

Bordes, J., R. Dumas de Vaulx, A. Lapierre, and M. Pollacsek. 1997. Haplodiploidization of maize (*Zea mays* L.) through induced gynogenesis assisted by glossy markers and its use in breeding. Agronomie 17:291–297.

Calus, M.P.L., T.H.E. Meuwissen, A.P.W. de Roos, and R.F. Veerkamp. 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178:553–561.

Chase, S.S. 1951. Production of homozygous diploids of maize from monoploids. Agron. J. 44:263–267.

Coe, E.H. 1959. A line of maize with high haploid frequency. Am. Nat. 93:381–382.

Concibido, V.C., B.W. Diers, and P.R. Arelli. 2004. A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci. 44:1121–1131.

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genomewide approach. PLoS ONE 3:E3395, DOI: 10.1371/journal.pone.0003395.

Eathington, S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47(S3):S154–S163.

Edwards, M., and L. Johnson. 1994. RFLPs for rapid recurrent selection. p. 33–40. *In* Analysis of molecular marker data. Joint Plant Breeding Symp. Ser., Am. Soc. Hortic. Sci., Corvallis, OR. 5–6 August 1994. CSSA, Madison, WI.

Gallais, A. 1990. Quantitative genetics of doubled haploid populations and application to the theory of line development. Genetics 124:199–206.

Gallais, A., and J. Bordes. 2007. The use of doubled haploids in recurrent selection and hybrid development in maize. Crop Sci. 47(S3):S190–S201.

Hallauer, A.R., and J.B. Miranda Fo. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames.

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

Hyne, V., M.J. Kearsey, D.J. Pike, and J.W. Snape. 1995. QTL analysis: Unreliability and bias in estimation procedures. Mol. Breed. 1:273–282.

Johnson, L. 2001. Marker assisted sweet corn breeding: A model for specialty crops. p. 25–30. *In* Proc. 56th Annu. Corn Sorghum Ind. Res. Conf., Chicago, IL. 5–7 Dec. 2001. Am. Seed Trade Assoc., Washington, DC.

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

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

Melchinger, A.E., C.F. Longin, H.F. Utz, and J.C. Reif. 2005. Hybrid maize breeding with double haploids lines: Quantitative genetic and selection theory for optimum allocation of resources. p. 8–21. Proc. 41st Annu. Illinois Corn Breeders School, Urbana, IL. 7–8 Mar. 2005. Univ. of Illinois at Urbana-Champaign.

Melchinger, A.E., H.F. Utz, and C.C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Moreau, L., A. Charcosset, F. Hospital, and A. Gallais. 1998. Marker-assisted selection efficiency in populations of finite size. Genetics 148:1353–1365.

Pumphrey, M.O., R. Bernardo, and J.A. Anderson. 2007. Validating the *Fhb1* QTL for Fusarium head blight resistance in near-isogenic wheat lines developed from breeding populations. Crop Sci. 47:200–206.

Röber, F., G.A. Gordillo, and H.H. Geiger. 2005. In vivo haploid induction in maize: Performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50:275–283.

Seitz, G. 2005. The use of doubled haploids in corn breeding. p. 1–7. *In* Proc. 41st Annu. Illinois Corn Breeders School, Urbana, IL. 7–8 Mar. 2005. Univ. of Illinois at Urbana-Champaign.

Senior, M.L., E.C.L. Chin, M. Lee, J.S.C. Smith, and C.W. Stuber. 1996. Simple sequence repeat markers developed from maize sequences found in the GENBANK database: Map construction. Crop Sci. 36:1676–1683.