Theor Appl Genet (2013) 126:2907–2920
DOI 10.1007/s00122-013-2182-x

ORIGINAL PAPER

Selfing for the design of genomic selection experiments in biparental plant populations

Benjamin McClosky · Jason LaCombe · Steven D. Tanksley

Received: 26 September 2012 / Accepted: 12 August 2013 / Published online: 27 August 2013

Abstract Self-fertilization (selfing) is commonly used for population development in plant breeding, and it is well established that selfing increases genetic variance between lines, thus increasing response to phenotypic selection. Furthermore, numerous studies have explored how selfing can be deployed to maximal benefit in the context of traditional plant breeding programs (Cornish in Heredity 65:201–211,1990a, Heredity 65:213–220,1990b; Liu et al. in Theor Appl Genet 109:370–376, 2004; Pooni and Jinks in Heredity 54:255–260, 1985). However, the impact of selfing on response to genomic selection has not been explored. In the current study we examined how selfing impacts the two key aspects of genomic selection—GEBV prediction (training) and selection response. We reach the following conclusions: (1) On average, selfing increases genomic selection gains by more than 70 %. (2) The gains in genomic selection response attributable to selfing hold over a wide range population sizes (100–500), heritabilities (0.2–0.8), and selection intensities (0.01–0.1). However, the benefits of selfing are dramatically reduced as the number of QTLs drops below 20. (3) The major cause of the improved response to genomic selection with selfing is through an increase in the occurrence of superior genotypes and not through improved GEBV predictions. While performance of the training population improves with selfing (especially with low heritability and small population sizes), the magnitude of these improvements is relatively small compared with improvements observed in the selection population. To illustrate the value of these insights, we propose a practical genomic selection scheme that substantially shortens the number of generations required to fully capture the benefits of selfing. Specifically, we provide simulation evidence that indicates the proposed scheme matches or exceeds the selection gains observed in advanced populations (i.e. F_8 and doubled haploid) across a broad range of heritability and QTL models. Without sacrificing selection gains, we also predict that fully inbred candidates for potential commercialization can be identified as early as the F_4 generation.

Communicated by F. van Eeuwijk.

B. McClosky (✉) · J. LaCombe · S. D. Tanksley
Nature Source Genetics, 33 Thornwood Drive Suite 300, Ithaca, NY 14850, USA
e-mail: bmcclosky@naturesourcegenetics.com

Introduction

In their seminal paper on genomic estimated breeding values (GEBVs), Meuwissen et al. (2001) claimed that the ability to predict genetic values directly from molecular marker data will have significant consequences for selection and breeding. Specifically, GEBV predictions have the potential to drastically reduce phenotyping costs and increase genetic gain in plants and animals by shortening generation intervals (Meuwissen et al. 2001). Subsequent empirical and simulation experiments support the notion that genomic, i.e. GEBV-based, selection will in fact revolutionize selection and breeding (Bernardo and Yu 2007; Goddard and Hayes 2007; Heffner et al. 2009; Legarra 2008; Lee et al. 2008; Lorenzana and Bernardo 2009; Luan et al. 2009; Schaeffer 2006).

The performance of genomic selection depends on two critical factors: (i) the occurrence of superior genotypes (i.e. transgressive segregants) in the selection population and (ii) the ability of GEBVs (based on a training population) to identify such individuals. When evaluating

and comparing experimental designs for genomic selection, gains in either factor can provide justification for the resource expenditure associated with population development. In this regard, plant researchers have at their disposal the option to advance populations through self-fertilization (or selfing). Selfing is a common mode of reproduction in plant species that affects multiple parameters relevant to genomic selection.

First, selfing increases homozygosity and amplifies additive genetic variance (Allard 1999; Falconer and Mackay 1996; Simmonds 1979). The resulting gains in additive genetic variance improve narrow-sense heritability for GEBV training and increase the occurrence of superior genotypes for selection. Second, selfing modulates the amount of effective recombination in a population (Falconer and Mackay 1996). Recombination reduces linkage disequilibrium (LD) among the markers, and reducing marker LD mitigates the negative impact of multicollinearity on marker-based regression models (Farrar and Glauber 1967). In addition, recombination can alter linkage phase among the QTL and thus have an impact on the breeding values of selection candidates (Iyamabo and Hayes 1995; Riggs and Snape 1977). Finally, populations undergoing progressive rounds of selfing are susceptible to genetic drift (Falconer and Mackay 1996; Wells and Weiser 1989), which can reduce genetic variance.

The purpose of this paper was to evaluate the use of progressive rounds of selfing in the design of genomic selection experiments. Researchers have previously addressed the question of which population characteristics affect the accuracy of popular GEBV training methods (Daetwyler et al. 2008, 2010; Zhong et al. 2009), and the impact of selfing on selection gains in traditional breeding has also been studied (Cornish 1990a, b; Liu et al. 2004; Pooni and Jinks 1985). However, the benefits of selfing for improving GEBV training and genomic selection response have not been considered in the literature. Such information can be important in deciding the number of selfing generations to employ during population development for GEBV training and/or selection.

In an effort to remedy this situation, we conducted a simulation study to quantify and contrast the performance of genomic selection in populations derived from a cross between two inbred lines. Such biparental populations are commonly developed as part of ongoing breeding programs and have been the focus of a number of recent research studies concerning genomic selection in plants (Bernardo 2010; Bernardo and Yu 2007; Guo et al. 2012; Heffner et al. 2010, 2011; Lorenzana and Bernardo 2009; Piepho 2009; Piyasatian et al. 2007; Wong and Bernardo 2008). Herein we investigate the impact of selfing on the response of biparental populations to genomic selection. Specifically, we address the following questions:

1. How does selfing impact the performance of genomic selection in terms of both training and selection response?
2. How many selfing generations are necessary to capture most of the performance benefits?
3. Does selfing the training population improve GEBV predictions when predicting outside of the training population?
4. Can the breeder capture gains comparable to those achieved in advanced populations (e.g. F_8 or doubled haploids) without investing the time and resources required to develop such advanced populations?

These results may be useful to researchers when analyzing the costs and benefits of selfing to maximize returns from genomic selection in breeding programs.

Methods

Simulation

We focused on populations derived from a biparental inbred cross through single seed descent (Allard 1999). We also included doubled haploid (DH) populations as a benchmark. These populations have high levels of linkage disequilibrium (LD) and, since they derive from only two inbred lines, contain only a small portion of the genetic variation present in the species (Snape 1976; Bordes et al. 2007; Bernardo 1996). Let P_1 and P_2 be the inbred parents used for population development.

Simulating the genome

We simulated the maize genome using the integrated map found in McMullen et al. (2009). All ten chromosomes had 1 cM marker spacing (1,350 markers in total). Crossovers were simulated according to a Poisson process using cM lengths derived from Haldane's mapping function. In each simulation iteration, we placed N_{QTL} QTLs randomly throughout the genome and sampled additive effects (a_1,…, a_{N_QTL}) from a gamma distribution with shape parameter 1.45, defined to be consistent with empirical evidence in maize (Buckler et al. 2009). A phase parameter p controlled the probability that a given effect a_j was negative. In other words, at a given QTL, p controlled the probability that the P_2 allele had a negative effect relative to the P_1 allele.

Simulating phenotype

Let $g_{ij} \in \{0,1,2\}$ encode the genotype of individual i at QTL j, and let $bv_i = \sum_{j=1}^{N_{QTL}} g_{ij}a_j$ denote individual i's breeding value. Unless otherwise stated, we simulated phenotype y according to the standard additive model: $y_i = bv_i + r_i$, where $r_i \sim N(0,\sigma_r^2)$. This model assumed constant environmental variance across breeding generations.

We also considered two classes of non-additive phenotypes. The first class transformed the additive breeding value to determine the genetic component of phenotype. Given a transformation $f : R \mapsto R$, we simulated phenotype using the following model: $y_i = f(bv_i) + r_i$. For instance, the sigmoid transformation in Fig. 1a modeled duplicate and complementary gene dosage (Kearsey and Sturley 1984). The parabolic transformation in Fig. 1b modeled biological systems where marginal increases in additive breeding value transition from being beneficial to detrimental. For example, maximizing the breeding value for plant height will increase the phenotype until the plant becomes too tall for the root system to support. The second class of non-additive model explicitly simulated pairwise epistatic interactions between alleles. We randomly assigned the number of interacting pairs uniformly between 1 and $\binom{N_Q}{2}$ and we sampled the epistatic effects from the same gamma distribution as the additive effects. The sign of each epistatic effect was set according to the phase parameter p.

Regardless of the model used to determine breeding value, we also introduced a normally distributed environmental noise term drawn from $N(0,\sigma_r^2)$. Let $\sigma_{F_2}^2$ be the variance of breeding values in an F_2. Let $h^2 = \sigma_{F_2}^2/\sigma_y{}^2$ denote heritability, and note that heritability always refers to F_2 heritability.

Parameter distributions

Each simulation iteration was parameterized by five key variables: population size (N), heritability (h^2), allele effect phase (p), number of QTL (N_{QTL}), and selection intensity (I). Selection intensity was defined as N_{sel}/N, where N_{sel} denotes the number of selected individuals. In order to study a broad range of scenarios, we sampled these variables from the default distributions shown in Table 1. The values for N, p, and I correspond to those commonly encountered in plant breeding. The range of h^2 (0.2–0.8) was chosen to represent values empirically observed in plant breeding experiments for quantitative traits (e.g. yield) (Hallauer and Miranda 1988; Albrecht et al. 2011; Bernardo 1996). The range of N_{QTL} (5–100) is consistent with empirical estimates of the number of QTLs segregating in experimental plant populations (Buckler et al. 2009; Laurie et al. 2004; Brachi et al. 2010; Otto and Jones 2000).

The non-additive models required additional parameterization. We varied the slope of the sigmoid transformation, $(1.0 + e^{-c_1 x})^{-1}$, by drawing c_1 uniformly from 0.1 to 1.0. For the parabolic transformation, we randomly sampled c_2^1, c_2^2 from the interval 0.0–1.0 and mapped the additive breeding values (via scaling) into the interval bounded by c_2^1 and c_2^2. This approach allowed us to use subsets of the parabolic transformation to model a broad range of non-additivity.

GEBV training methods

For the majority of our experiments, we used GBLUP to compute GEBVs (Meuwissen et al. 2001). To demonstrate that our results are robust to the choice of estimation methodology, we also examined the impact of using Bayes B

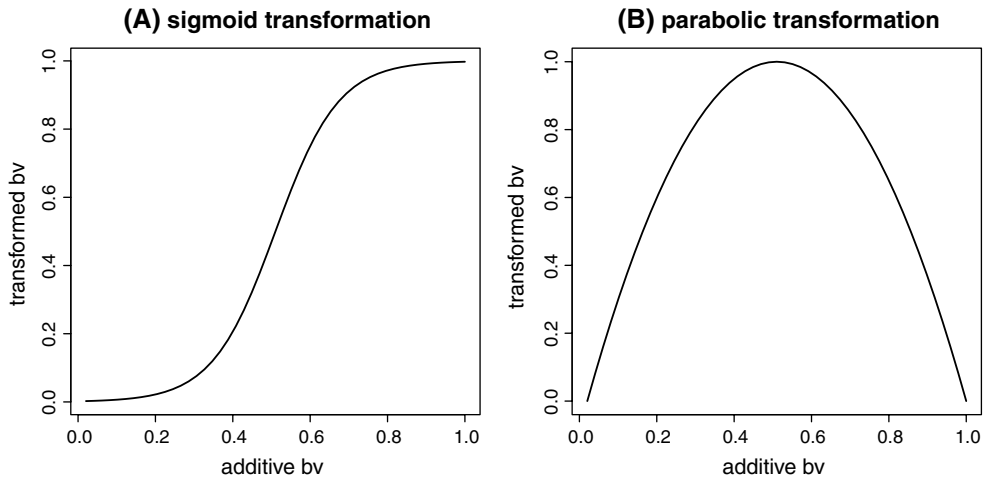

Fig. 1 Examples of non-additive relationships between additive breeding value and the transformed breeding value used to determine phenotype. Simulated breeding values were scaled into the [0,1] interval and mapped using the transformation: (**a**) $f(x) = (1.0 + e^{-c_1 x})^{-1}$ or (**b**) $f(x) = 1 - 4(x - 0.5)^2$

Table 1 Default distributions for simulation parameters. Parameters were randomly sampled within the specified range for each iteration of the simulation

Parameters	Symbols	Range
Number of QTL	N_{QTL}	5–100
Heritability	h^2	0.2–0.8
Phase	p	0.4–0.6
Selection intensity	I	0.01–0.1
Population size	N	100–500

and Bayes C to compute GEBVs (Meuwissen et al. 2001; Verbyla et al. 2009, 2010).

The Bayes B method incorporates variable selection into the estimation of GEBVs through the specification of a parameter π, which represents the prior probability of a variable having a nonzero effect. Like GBLUP, Bayes B assumes that the (non-zero) effects are normally distributed, though with effect-specific variance components. The GEBVs were obtained by the technique of Bayesian Model Averaging (Hoeting et al. 1999) across samples obtained by the application of the Metropolis–Hastings Algorithm (Hastings 1970).

The method of Bayes C is similar to that of Bayes B, the difference being that the parameter π is estimated from the data or assumed to be distributed according to some prior distribution. The estimation of π allows Bayes C the flexibility to adjust the proportion of nonzero effects for a given dataset. In our implementation of Bayes C, we estimate π as a function of the observed genotypes and map data.

Performance metric for response to genomic selection

Let σ_{F_2} be the standard deviation of breeding values in the F_2. To simulate genomic selection, we sorted the selection population by GEBV, selected the top N_{sel} individuals, and determined the maximum true breeding value in the selected set. Gain was reported as

$$(\max_{i \in N_{sel}}\{bv_i\} - \max\{bv_{P_1}, bv_{P_2}\})/\sigma_{F_2}.$$

In other words, we measured advancement of the maximum selected individual beyond the best parent. This transgressive gain metric reflects our interest in identifying the best possible individual for inbred development. In this context, we use the term *superior genotypes* to refer to the most desirable candidates for selection, i.e. the individuals with the largest breeding values. Note that gain was always defined in units of standard deviation of F_2 breeding values.

Selection populations

For the first set of experiments (Sects. 3.1–3.2), we performed genomic selection in the same population used for GEBV training, i.e. the training and selection populations coincided. For the second set of experiments (Sects. 3.3–3.4), we separated the training and selection populations. More precisely, the training and selection populations were both derived from the same biparental family, but we simulated the populations independently. Finally, for the third set of experiments (Sects. 3.6–3.7), we performed genomic selection in the progeny of high-performing training set individuals. We designed this last set of experiments to evaluate a new method for shortening the number of selection and inbreeding cycles required to produce fully inbred candidates for potential commercialization.

Results

Impact of selfing on gains when training and selecting from the same population

Figure 2a clearly demonstrates that progressive selfing had a significant and positive impact on genomic selection gains. In particular, selfing to the F_8 produced a 72 % increase over F_2 gains. Figure 2a also shows that marginal gains diminished as selfing progressed. For example, the F_4 captured nearly 90 % of the F_8 gains, and the F_5 realized 95 % of the F_8 gains. Simulation has led to similar

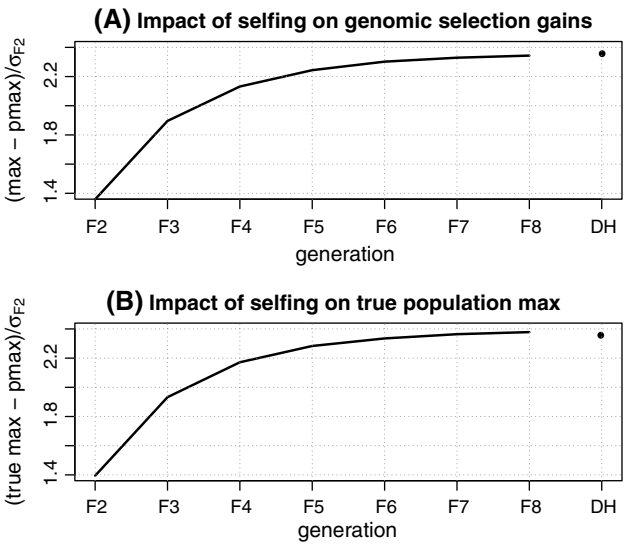

Fig. 2 The effect of selfing on (**a**) genomic selection gains and (**b**) population max true breeding value. On the *y*-axes, *max* denotes the best true breeding value individual in the N_{sel} subset, *true max* denotes the best true breeding value in the population (regardless of whether it was obtained in N_{sel}), and *pmax* denotes the best parental true breeding value. This figure shows average performance across the parameter distributions listed in Table 1. Based on 10,000 simulations (S.E. $\leq$ 0.01)

conclusions in the context of phenotypic selection (Cornish 1990). Also note that the F_8 and DH performed similarly, consistent with previous observations (Bordes et al. 2007; Murigneux et al. 1993; Park et al. 1976; Choo et al. 1982; Courtois 1993).

As noted earlier, response to genomic selection depends both on the quality of GEBV training and the magnitude of the max breeding value in the selection population. These two factors are confounded in Fig. 2a, which depicts the max breeding value captured by genomic selection. As a result, the specific cause of improved response to genomic selection upon selfing cannot be determined. In an effort to resolve this issue, we generated a second plot showing the change in the true max population breeding value (regardless of whether it was captured by genomic selection) across selfing generations (Fig. 2b). The similarity of the curves in Fig. 2a, b suggests that the majority of gains due to selfing arise from an increase in the population max breeding value, rather than improvement in GEBV training. This assertion is also supported by experiments (described in a later section) which use independent training and selection populations. However, we first wish to quantify the marginal effects of the critical simulation parameters when training and selecting from the same population.

Marginal impact of simulation parameters when training and selecting from the same population

The curve in Fig. 2a reports the expected genomic selection gain in each selfing generation from the F_2 to the F_8. Notice that the maximum expected gain occurs in the F_8 and the minimum expected gain occurs in the F_2. For the experiments in this section, we defined Δgain as the difference between the max gain and min gain observed in the sequence of selfing generations (F_2–F_8). The Δgain metric quantifies the impact that repeated selfing had on selection gains. Positive values for Δgain indicate that selfing affected genomic selection gains.

Figure 3 shows how Δgain responded to changes in the primary simulation parameters. Population size (N) had very little impact on Δgain once $N \geq 200$, though Δgain increased as N fell to 100 (Fig. 3a). Similarly, changes in heritability had only a minor influence on Δgain once $h^2 \geq 0.5$, but Δgain increased modestly when heritability decreased below this value (Fig. 3b). The magnitude of Δgain was sensitive to selection intensity (I) when $0.01 \leq I \leq 0.05$, but the influence of selection intensity diminished as I exceeded 0.05 (Fig. 3c). The number of QTL (N_{QTL}) clearly had the greatest impact on Δgain (Fig. 3d). In particular, Δgain increased very rapidly as N_{QTL} increased from 5 to 25 but had less response thereafter. In

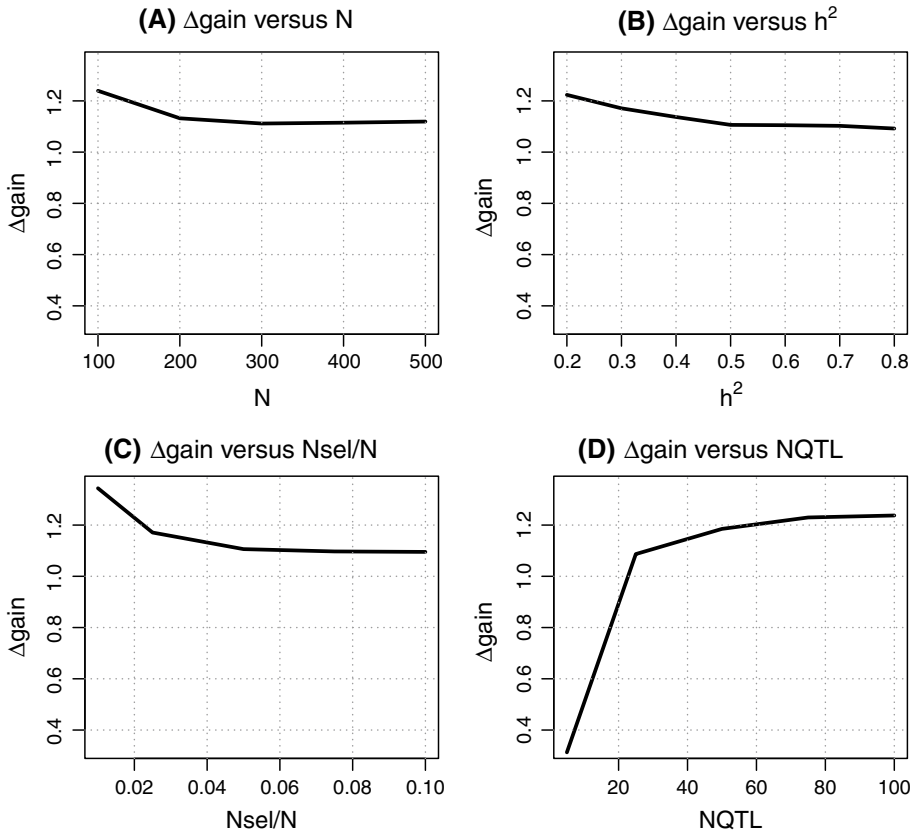

Fig. 3 The impact of the key simulation parameters on Δgain when training and selecting in the same population. Δgain measures the impact that selfing had on genomic selection gains (i.e. Δgain = max gain − min gain). Positive values for Δgain indicate that selfing affected genomic selection gains. These curves illustrate how Δgain responded to changes in (**a**) population size; (**b**) heritability; (**c**) selection intensity; (**d**) number of QTL. All parameters not specified by the *x*-axis were drawn from default distributions in Table 1. Based on 5,000 simulations (S.E. ≤ 0.01)

summary, the curves stayed strictly above 0.0 across all values for the critical variables, indicating that selfing consistently had an impact on genomic selection gains.

The Δgain metric carried insufficient information to determine how many rounds of selfing were necessary to capture the majority of potential gains. To clarify this issue, we defined G to be the earliest generation such that gains exceeded 90 % of the maximum gain achievable through selfing. By tracking G, we could determine the minimum number of generations necessary to realize most of the benefits from selfing. Figure 4 illustrates how the main simulation parameters affected G.

In most cases, selfing to the F_5 was sufficient to capture 90 % of the potential gains (Fig. 4). This conclusion was largely independent of population size (Fig. 4a), heritability (Fig. 4b), and selection intensity (Fig. 4c). However, the number of QTL had a significant impact on G (Fig. 4d). As the number of QTLs dropped below 10, selfing to the F_3 sufficed to capture most of the potential gains.

Impact of selfing on gains when training and selecting in independent populations

As previously stated, when training and selecting in the same population, we cannot determine whether selfing improved genomic selection response through better GEBV training or through the increased occurrence of superior genotypes. While Fig. 2 indicates that selfing improved selection gains by increasing the max population breeding value, it remains to determine whether selfing also significantly improved the ability of the GEBVs to identify superior genotypes.

To study the impact of selfing on training, we trained GEBVs on selfed populations ($F_2 - F_8$) and performed selection in different DH and F_8 populations. The selection populations were independently derived from the same biparental cross. Figure 5 shows the impact of selfing the GEBV training population when selecting in (A) the DH and (B) the F_8. Regardless of whether selecting from a DH or F_8, selfing produced only minor improvement in genomic selection gains. This evidence supports the hypothesis that selfing the training population only modestly improves the ability of the GEBVs to identify superior genotypes. This observation is consistent with our previous experiments where training and selection occurred in the same population. In that case, an increase in the population max appeared to be the dominant factor in determining selection gains (Fig. 2).

It should be noted that these results correspond to average performance across a range of values for the key simulation parameters (Table 1). In the next section, we test the sensitivity of our conclusions to marginal changes in each

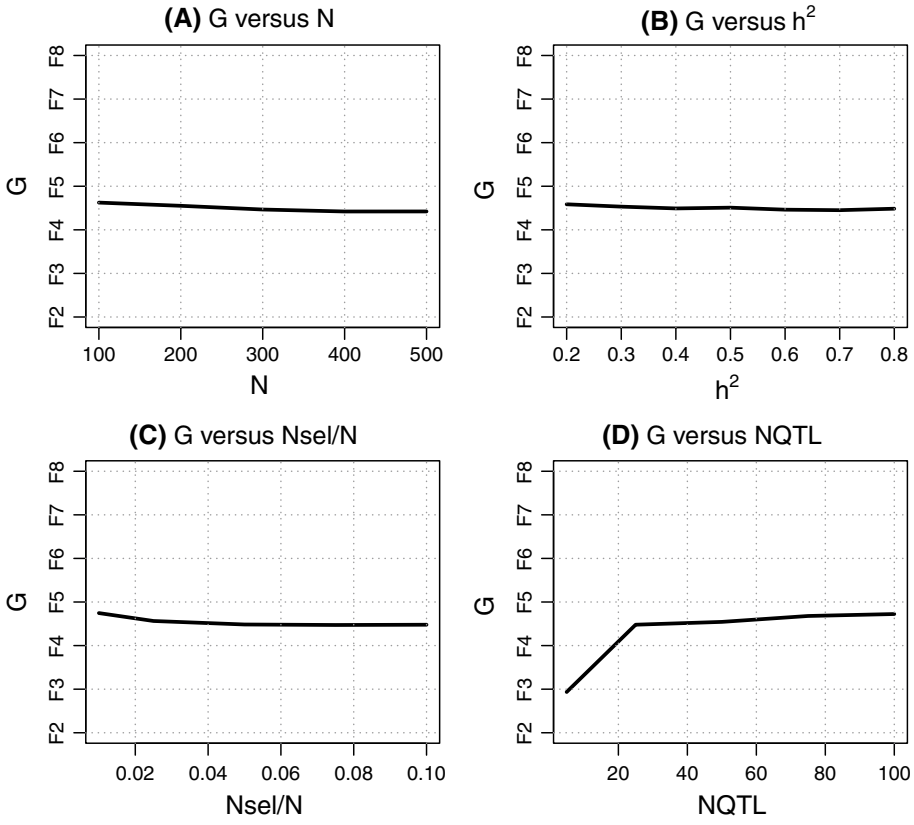

Fig. 4 The impact of the key simulation parameters on G when training and selecting in the same population. G denotes the earliest generation where gains exceeded 90 % of the maximum gain achievable through selfing. These curves show how G responded to changes in (**a**) population size; (**b**) heritability; (**c**) selection intensity; (**d**) number of QTL. All parameters not specified by the x-axis were drawn from default distributions in Table 1. Based on 5,000 simulations

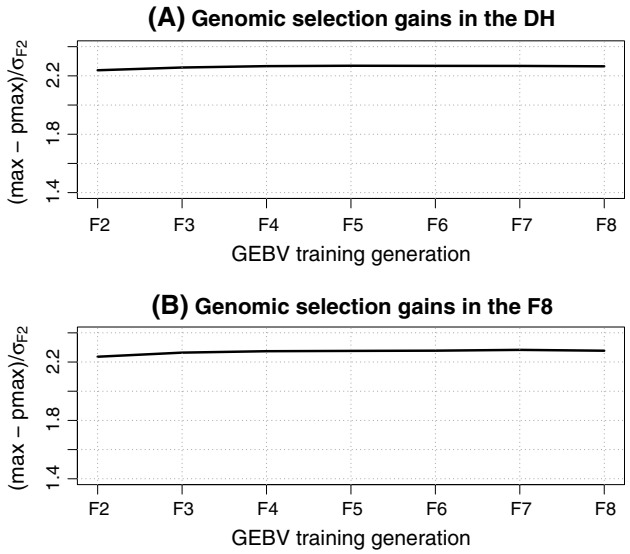

Fig. 5 Impact of selfing on GEBV training when selecting in (**a**) the DH and (**b**) the F_8. The *x*-axis reports the GEBV training generation. Training and selection populations were bred independently from the same biparental cross. On the *y*-axis, *max* denotes the best true breeding value in the N_{sel} subset and *pmax* denotes the best parental true breeding value. This figure shows average performance across the distributions in Table 1. Based on 10,000 simulations (S.E. ≤ 0.01)

key parameter value. Further, since the DH and F_8 selection populations gave nearly identical results in this section, experiments in the next section use only the DH population for selection.

Marginal impact of simulation parameters when training and selecting in independent populations

The results in the previous section indicate that, on average, selfing the training population does not significantly impact genomic selection response in an independent selection population. In this section, we describe experiments conducted to determine the sensitivity of this conclusion to marginal changes in the key simulation parameters (Table 1). For these experiments, we defined Δgain as the difference between the max gain and min gain observed (selecting in the DH) in the sequence of training generations (F_2–F_8). We then determined how the primary simulation parameters affected Δgain. The Δgain metric quantifies the impact that repeated selfing of the training set had on selection gains. Positive values for Δgain indicate selfing impacted gains.

A general observation is that Δgain was substantially lower when training and selecting in independent populations (Fig. 6) than when training and selecting in the same population (Fig. 3). This is despite the fact that the training and selection populations derived from the same biparental cross and had similar allele frequencies. In fact, Δgain approached 0.0 (indicating no benefit from selfing) in several cases: $N = 500$ (Fig. 6a), $h^2 = 0.8$ (Fig. 6b), and $N_{sel}/N = 0.1$ (Fig. 6c). N_{QTL} had the least impact on Δgain (Fig. 6d). This observation is consistent with previous claims that N_{QTL} has little impact on GBLUP performance (Daetwyler et al. 2010). The Δgain metric became responsive to the marginal parameters as N approached 100, h^2 approached 0.2, or Nsel/N approached 0.01 (Fig. 6a–c). Under these conditions, training significantly improves with selfing. However, even in these cases, selfing the training set had less effect on gains than selfing the selection population (Fig. 3). Overall, this evidence supports the hypothesis that the positive impact of selfing on genomic selection is largely due to increases in the occurrence of superior genotypes rather than improvement in GEBV training.

As noted previously, Δgain carried insufficient for determining how many selfing generations were necessary to accrue most of the performance benefits. Thus, we again employed the G statistic, which is defined as the earliest training generation exceeding 90 % of the maximum gain observed (Fig. 7). As seen in Fig. 4, selfing to the F_5 was generally sufficient to capture the majority of potential gains, but the F_3 became sufficient as the number of QTLs dropped below ten. This evidence suggests that GEBV gains from selfing experience diminishing marginal returns.

Summary of key issues regarding the impact of selfing on genomic selection

Thus far the simulation evidence supports the following conclusions:

1. Selfing of the selection population significantly improves response to genomic selection by increasing the occurrence of superior genotypes.
2. Selfing of the training population can increase response to genomic selection by improving the ability of the GEBVs to identify superior genotypes. However, this benefit is marginal compared with the benefits of selfing the selection population.
3. Selfing to the F_5 captures, on average, more than 90 % of the gains observed in the F_8 population.

The first two conclusions suggest that selfing for genomic selection is best applied to increase the frequency of transgressive segregants, rather than to improve GEBV training. The third conclusion indicates that it might be possible to reduce the number of generations required to realize the benefits of selfing for genomic selection response.

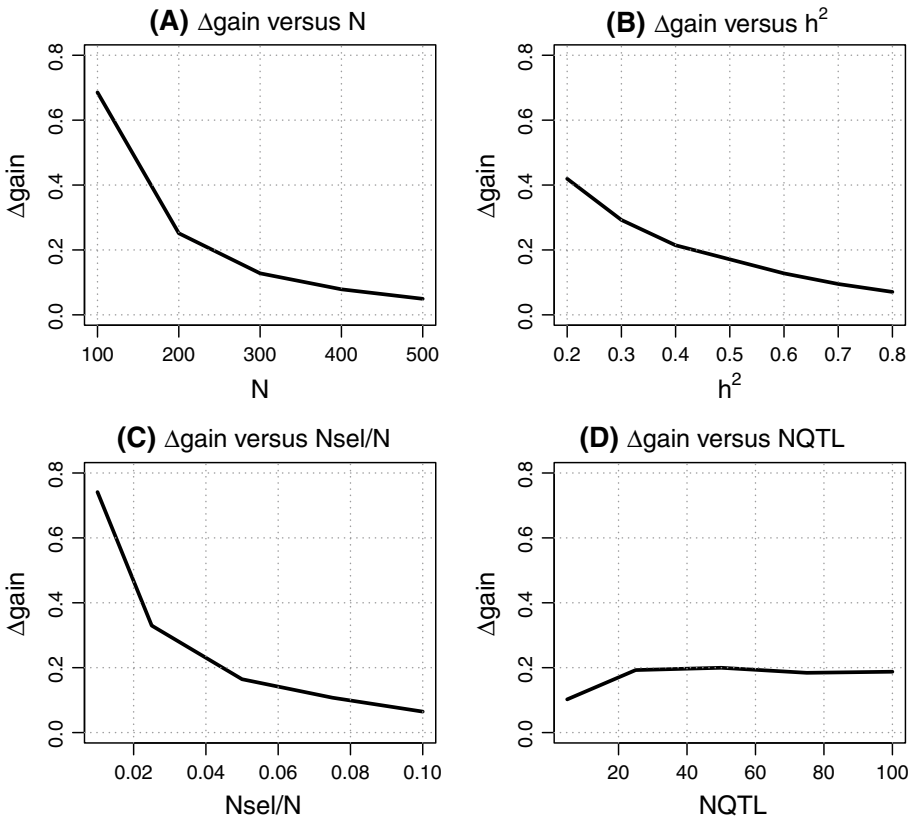

Fig. 6 The impact of the key simulation parameters on Δgain when training and selecting in independent populations. Δgain measures the impact that selfing had on genomic selection gains (i.e. Δgain = max gain – min gain). Positive values for Δgain indicate that selfing the training set affected genomic selection gains. These curves illustrate how Δgain responded to changes in (**a**) population size; (**b**) heritability; (**c**) selection intensity; (**d**) number of QTL. All parameters not specified by the *x*-axis were drawn from default distributions in Table 1. Based on 5,000 simulations (S.E. ≤ 0.01)

Is it possible to reduce the number of generations while still realizing the benefits of selfing for response to genomic selection?

The above results suggest that it is not necessary to create advanced selfed generations to obtain good training for genomic selection. The main benefit of selfing lies in the production of superior genotypes for selection. This led to the question of whether one might train in an early generation (e.g. F_2, F_2 testcross, or F_3 families – depending on the crop) and use the resulting GEBVs to select from a larger population of segregating individuals for which only genotype information was available. Specifically, we would like to construct a selection population with properties as good or better than training and selection in an advanced selfed generation (e.g. F_8), but we want to avoid investing the time and resources required to produce such advanced generations. As the price of genotyping continues to fall, we expect marker screening large populations in early generations to become cost effective.

During progressive rounds of selfing, additive genetic variance within the lines vanishes as additive genetic variance between the lines increases (Lynch and Walsh 1998; Cornish 1990). Therefore, the previously observed F_8 genomic selection gains were a consequence of selecting on additive genetic variance between the lines. In order to reduce the number of selfing generations required for development of the selection population, we employed genomic selection to match or exceed the F_8 gains by selecting on F_2 within line variance (i.e. the F_2 progeny). The resulting genomic selection scheme (outlined in Fig. 8) substantially reduces the time and selfing generations required to develop inbreds for potential commercialization.

In the remainder of this section, we describe simulation experiments conducted to evaluate the scheme in Fig. 8. We first simulated 200 F_2 individuals with phenotypes for GEBV training. It should be noted here that the phenotype of the F_2 may refer to any of the following depending on the crop: (1) The phenotype of the actual F_2 individual. This would apply to the small number of situations in which single plant phenotyping is valid, such as for high heritability traits in glasshouse grown vegetable crops or perennial tree species; (2) F_3 or F_4 plots derived from each F_2 individual; (3) Testcross plots derived from crossing each F_2 individual to one or more tester lines. Using a training population of size 200 is consistent with previous experiments exploring the application of genomic selection in biparental crosses (Bernardo and Yu 2007).

GEBVs were then used to select the top 20 F_2 individuals ($I = 0.1$ based on Figs. 3c, 6c), and an equal number of selfed progeny were drawn from each of the selected F_2 individuals. The resulting F_3 population was genotyped

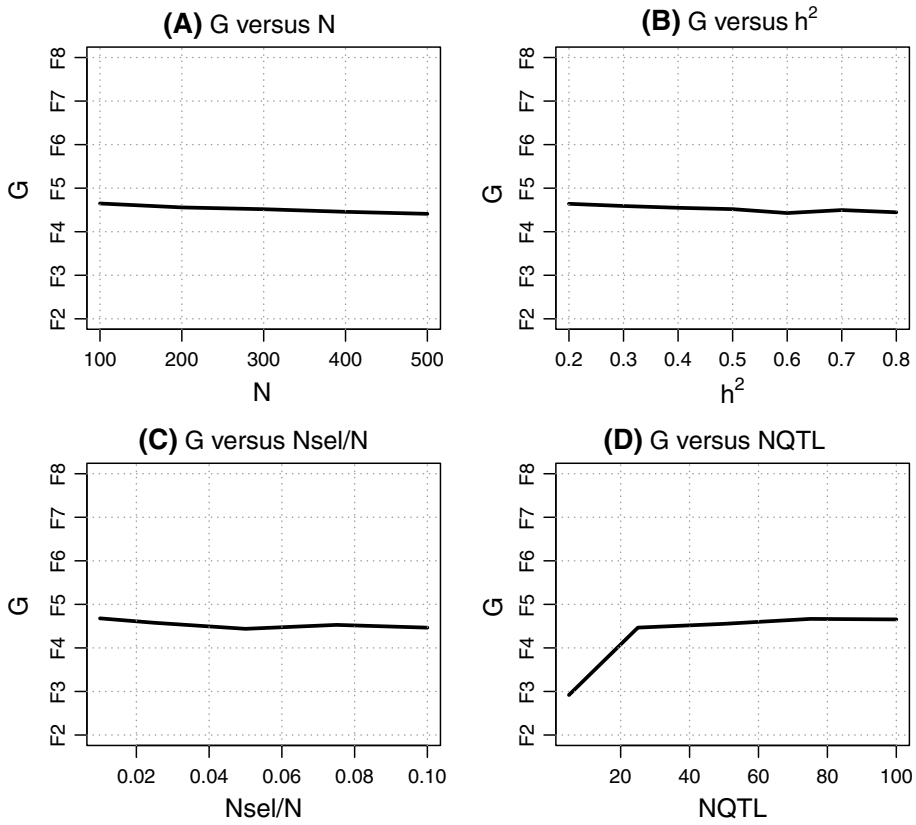

Fig. 7 The impact of the key simulation parameters on *G* when training and selecting in independent populations. *G* denotes the earliest generation where gains exceeded 90 % of max gains observed when selecting in the *DH* and training in the $F_2 - F_8$. These curves show how *G* responded to changes in (**a**) population size; (**b**) heritability; (**c**) selection intensity; (**d**) number of QTL. All parameters not specified by the *x*-axis were drawn from default distributions in Table 1. Based on 5,000 simulations

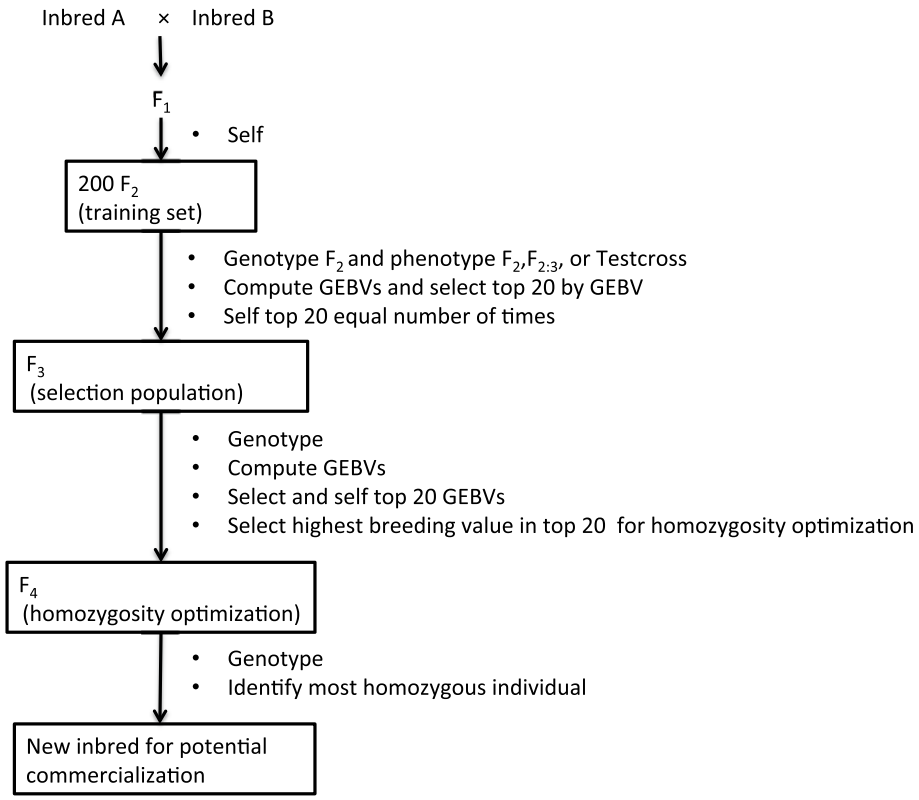

Fig. 8 Genomic selection breeding scheme designed to reduce time and selfing generations required for inbred development

and sorted by GEBVs, and the highest true breeding value was recorded among the top 20 F_3 individuals. In simulation, we can simply observe the true breeding values. In practice, breeders would send the top 20 GEBV plants to variety trials to obtain estimates of true breeding values. Finally, we sampled and genotyped progeny from the max selected F_3 individual to optimize homozygosity for inbred development.

The first objective was to determine how many F_3 individuals to screen to match or exceed gains obtained by training and selection in the F_8. According to Fig. 9, an F_3 of size 300 (i.e. sampling 15 selfed progeny from each of the top 20 F_2 individuals) matched the F_8 gains, and the F_3 significantly outperformed the F_8 thereafter. By sampling 1,500 F_3 individuals, we increased average performance by more than 10 % over the F_8 gains. Figure 9 also illustrates diminishing marginal returns as the F_3 population size grows past 1,500.

The next goal was to determine how F_3 gains compared with F_8 gains as we varied the key simulation parameters. Figure 10 shows how both schemes performed across a range of marginal values for heritability and number of QTL. This experiment fixed the F_3 size at 1,500 plants (i.e. 75 seed per selected F_2 individual). Figure 10a indicates that the schemes captured comparable gains at low heritability, but the F_3 scheme clearly dominated as h^2 increased.

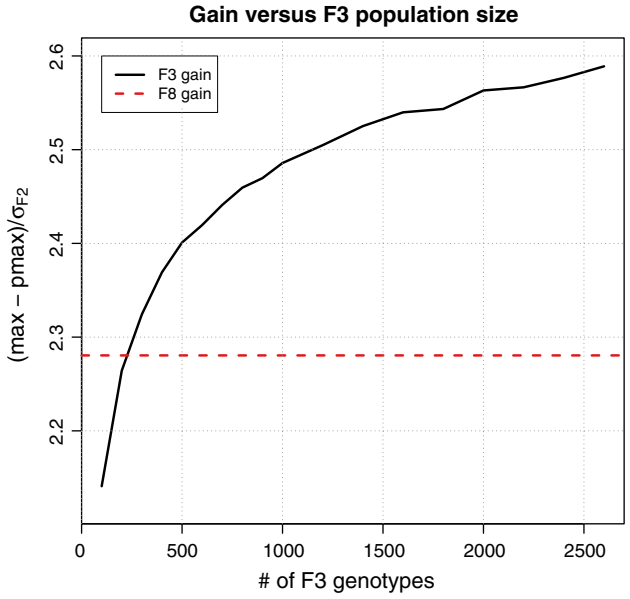

Fig. 9 Effect of F_3 population size on gains from genomic selection. On the *y*-axis, *max* denotes the best true breeding value in the selected subset of F_3 individuals ($N_{sel} = 20$) and *pmax* denotes the best parental true breeding value. The *red dashed line* shows genomic selection gains when training and selecting in a population of 200 F_8 individuals ($N_{sel} = 20$). This figure shows average performance across the distributions in Table 1. Based on 100,000 simulations (S.E. ≤ 0.01)

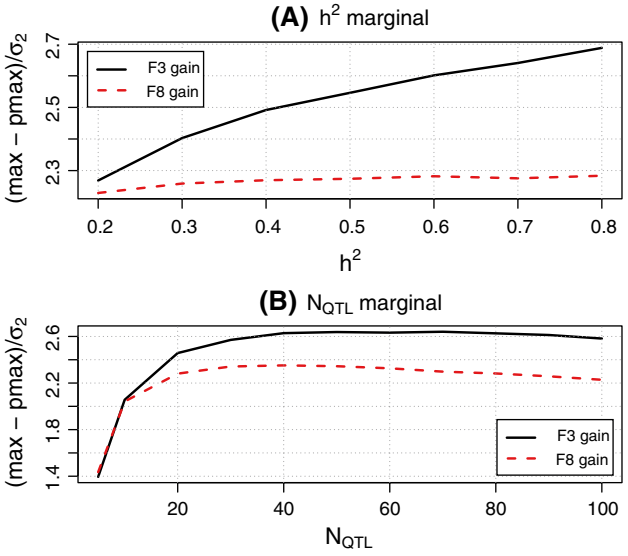

Fig. 10 Comparison of F_3 and F_8 genomic selection gains while varying key simulation parameters: (**a**) heritability and (**b**) number of QTL. On the *y*-axis, *max* denotes the best true breeding value in the selected subset of F_3 individuals ($N_{sel} = 20$) and *pmax* denotes the best true parental breeding value. The *red dashed line* shows genomic selection gains when training and selecting in a population of 200 F_8 individuals ($N_{sel} = 20$). For the F_3 scheme, F_3 population size was set to 1,500 (i.e. 75 seed per selected F_2 individual). All parameters (except N, N_{sel} and the *x*-axis) were drawn from default distributions in Table 1. Based on 100,000 simulations (S.E. ≤ 0.01)

Overall, heritability had significantly more impact on F_3 selection than F_8 selection. Figure 10b shows comparable gains under sparse QTL models, but the F_3 scheme again dominated as N_{QTL} increased.

Although F_3 selections surpassed the F_8 gains, thereby saving several generations, there was a price to pay in heterozygosity. For the F_3 scheme, the selected max individual averaged 22 % heterozygosity (not shown). Such high levels of heterozygosity are normally unacceptable for commercial production—either as inbreds for hybrid production or as open-pollinated varieties.

We thus asked the question whether an additional generation of selfing (F_4), combined with marker-assisted selection for homozygosity, might result in lines with acceptable levels of heterozygosity. For this experiment, we kept the F_3 size fixed at 1,500 as previously described. After performing genomic selection in the F_3, we selfed the max selected F_3 individual multiple times and reported the minimum heterozygosity observed in the progeny as a function of F_4 population size (Fig. 11). Since no trait selection was performed in the F_4, the F_4 selected individual had the same expected breeding value as the F_3 selected individual (not shown). In other words, the homozygosity optimization did not impact expected gains.

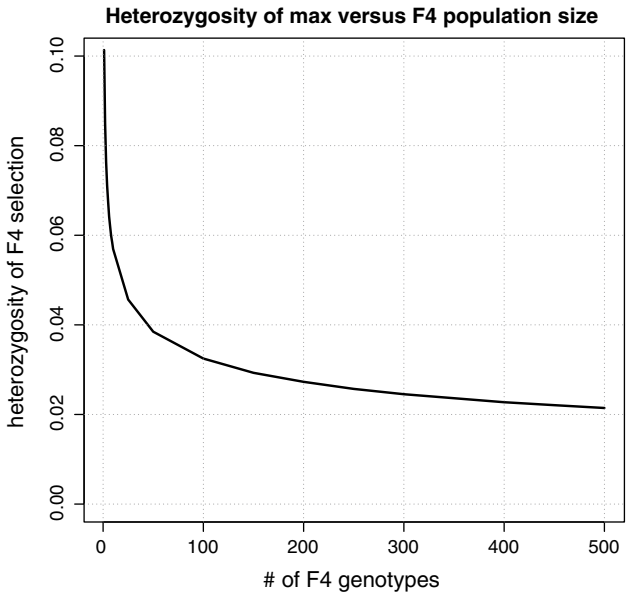

Fig. 11 The relationship between heterozygosity and F_4 population size when using markers to select against heterozygosity. The *y*-axis reports the minimum heterozygosity observed in the F_4 progeny. Based on 100,000 simulations (S.E. ≤ 0.001)

To summarize, the curves in Figs. 9, 10 indicate that—by screening 1,500 F_3 genotypes—we can match or exceed F_8 gains across a broad range of heritability and number of QTLs. In addition, without reducing expected gains, we can achieve a heterozygosity of about 0.03 by sampling 150 F_4 progeny and of about 0.02 by sampling 500 progeny (Fig. 11). Compared with the F_8 single seed descent approach, this scheme would save four generations of breeding and incur the cost of genotyping $\leq$2,000 additional individuals (1,500 F_3 and 500 F_4). Further, the new approach requires no additional phenotyping (Fig. 8).

The value of the proposed method is based on two primary assumptions. First, we have assumed that the cost of genotyping the F_3 and F_4 individuals does not greatly exceed the cost of developing an F_8 population. Second, we have assumed that sampling thousands of progeny is biologically feasible.

Impact of non-additivity and the GEBV training method on gains

The F_3 gains reported above were based on the assumption of additive effects and on the use of a single GEBV training method (GBLUP). In this section, we evaluate the impact of deviations from additivity and of using alternative GEBV models. Table 2 reports the performance of the F_3 scheme using three GEBV methods across three non-additive genetic models. For this experiment, we fixed the F_3 size at 1,500 plants (i.e. 75 seed per selected F_2 individual) and $N_{sel} = 20$. The first column (ADD) shows gains under additive phenotypes. The second column (NA1) shows gains in the presence of pairwise epistasis. The third (NA2) and fourth (NA3) column show performance under the sigmoid and parabolic transformations, respectively.

Table 2 The effects of GEBV method and non-additive effects models (NA) on F_3 genomic selection gains

Methods	ADD	NA1	NA2	NA3
BLUP	2.53 (0.01)	1.33 (0.01)	2.31 (0.01)	2.33 (0.01)
Bayes B	2.40 (0.01)	1.20 (0.03)	2.28 (0.01)	2.28 (0.01)
Bayes C	2.52 (0.01)	1.25 (0.03)	2.30 (0.02)	2.31 (0.01)

ADD denotes additive effects, and *NA1* denotes pairwise epistasis. NA2 used the sigmoid transformation of breeding value. NA3 used the parabolic transformation (see "Methods" for details). F_3 population size was fixed at 1,500 plants based on 10,000 simulations. Standard errors in parentheses

Table 3 Relative F_3 gains using different GEBV methods and non-additive effects models

Methods	ADD	NA1	NA2	NA3
BLUP	1.11	1.48	1.02	1.02
Bayes B	1.05	1.33	1.01	1.00
Bayes C	1.11	1.38	1.02	1.01

Note: This table expresses the results from Table 2 in terms of relative F_8 performance (F_3 genomic selection gain)/(F_8 max possible gain)

For all three training methods, non-additivity reduced genomic selection performance (Table 2). Pairwise epistasis (NA1) was most severe, reducing gains by about 50 %. With regards to GEBV methods, Bayes B was inferior over all models. Furthermore, given the similar performance of GBLUP and Bayes C, these results suggest that the use of Bayesian variable selection is unlikely to improve the performance of this particular F_3 scheme.

While Table 2 clearly demonstrates that non-additivity reduced F_3 genomic selection gains, we have not resolved the question of whether the F_3 scheme captured gains comparable to the F_8 scheme under non-additivity. Therefore, we computed a ratio of F_3 genomic selection gains to the best possible F_8 gains. Specifically, for each model of non-additivity, we computed the expectation of the max true breeding value in the F_8. This value bounds the performance of F_8 genomic selection gains and was independent of GEBV training method. Table 3 expresses the results from Table 2 as a proportion of the best possible gain when training and selecting in the F_8. In other words, Table 3 reports a lower bound on the relative performance of the F_3 scheme to the F_8 scheme.

This analysis leads to several interesting observations. First, pairwise epistasis (NA1) caused the F_3 scheme to have an even higher performance relative to the F_8.

In other words, the F_3 scheme was less susceptible to the negative effects of pairwise epistasis. We speculate that, by allowing small differences in genotype to produce major differences in breeding value, pairwise epistasis amplified the benefits of the large sample size (1,500 individuals) used in the F_3 scheme. Second, while F_3 gains exceeded the F_8 by more than 10 % under additivity (ADD), the F_3 performed approximately the same as the F_8 with the non-linear transformations (NA2, NA3). In summary, although non-additivity severely reduced absolute gains, the F_3 scheme (using BLUP or Bayes C) performed as well or better than the F_8 scheme under all types of non-additivity considered.

Discussion

Unlike animals, most plants are capable of self-fertilization, a severe form of inbreeding. Selfing not only amplifies additive genetic variance (i.e. heritability), and hence response to phenotypic selection, but it also leads to more uniform genetic stocks, which is critical in the development of both hybrid and open pollinated crops (Simmonds 1979). As a result, selfing is a common step in the breeding paradigm of most crop plants (Fess et al. 2011).

Genomic selection is a relatively new approach for maximizing selection gains in plant and animal breeding. Compared with traditional phenotypic selection, genomic selection has the potential to improve the precision of selection, increase the number of breeding cycles per unit time, and reduce phenotyping costs (Meuwissen et al. 2001). A successful genomic selection experiment requires the development of populations to serve two distinct but complementary purposes. First, the training population must have genotype data, phenotype data, and favorable properties for marker-based regression models. For example, marker density, LD, and heritability can all impact GEBV training (Zhong et al. 2009; Daetwyler et al. 2010). Second, the selection population—which requires genotype data but not necessarily phenotype data—must contain superior genotypes (i.e. transgressive segregants) that can be identified and selected for advancement using GEBVs. An important feature of genomic selection is that the training and selection populations need not coincide.

A number of studies on the application of genomic selection in plants have focused on situations where both the training and selection populations trace back to the same biparental cross(es), and selfing—or other severe forms of inbreeding such as doubled haploidy—has often been used to generate these populations (Mayor and Bernardo 2009; Bernardo and Yu 2007; Guo et al. 2012; Heffner et al. 2010, 2011; Piepho 2009). However, to our knowledge, none of these studies have examined the impact of selfing on response to genomic selection, nor have they separated the impact of selfing on the training population versus the selection population. Since selfing is a common step in plant breeding, it seemed worthwhile to investigate its impact on genomic selection.

Impact of selfing on genomic selection gains

The purpose of the current study was to quantify the impact of selfing on the performance of genomic selection in biparental populations over a wide range of key parameters (e.g. heritability, number of QTLs, population size and selection intensity). Moreover, we wished to measure the relative impact of selfing on the performance of the training population versus the selection population. Key findings relative to these issues are as follows:

1. Selfing has a significant positive impact on response to genomic selection. For example, F_8 gains exceeded F_2 gains by more than 70 % on average. However, most of the gains are realized by the F_5 generation.
2. The gains in genomic selection response attributable to selfing hold over a wide range population sizes (100–500), heritabilities (0.2–0.8), and selection intensities (0.01–0.1). However, the benefits of selfing are dramatically reduced as the number of QTLs drops below 20.
3. The major cause of improved response of genomic selection with selfing is through an increase in occurrence of superior genotypes (i.e. transgressive segregants) in the selection population.
4. Performance of the training population improves with selfing (especially with low heritability and small population sizes); however, the magnitude of these improvements is relatively small compared with improvements observed in the selection population.

Implications for practical application of genomic selection in plant breeding

The finding that training populations and selection population respond differently to selfing raises the question of how best to use selfing in breeding experiments involving genomic selection. Results from the current study indicate that is not necessary to create advanced selfing populations to achieve good training for genomic selection. In fact, across a wide range of key parameters (e.g. heritability, population sizes, number of QTLs), good training can be achieved as early as the F_2 (e.g. genotyping F_2 individuals then phenotyping either F_2 individuals, testcrosses or F_3 progeny). However, further selfing does significantly improve performance of the selection population, and thus genomic selection gains, by increasing the occurrence of superior genotypes.

The most expensive part of genomic selection experiments is the time and cost of phenotyping the training population. However, once training is accomplished, individuals in the selection population have only to be genotyped to derive GEBVs for selection. Thus, as part of the current study, we asked whether one could train in an F_2 population and then use that information to screen a larger F_3 population (via GEBVs) such that genetic gains that match or exceed an F_8 or DH might be achieved. The results from these experiments were encouraging, indicating that gains surpassing an F_8 or DH are achievable when selecting from an F_3 under a wide range of key parameters. One potential drawback of this method is the relatively high level of heterozyosity of lines selected at this stage. However, simulations indicated that marker-based selection can be used to reduce heterozygosity to <3 % by selfing the GEBV selected F_3 individuals.

Finally, it should be noted that the above was a proof-of-concept experiment using simulation, and that the specific strategy outlined would need to be modified and tailored to the biological constraints of any given crop. Nonetheless, these results are encouraging and suggest that optimal deployment of genomic selection in biparental populations might involve only a couple of rounds of selfing, while delivering gains traditionally achieved only after significantly more rounds of selfing (or DH).

Conflict of interest The authors declare no conflict of interest.

References

Albrecht T, Wimmer V, Auinger H, Erbe M, Knaak C, Ouzunova M et al. (2011) Genome-based prediction of testcross values in maize. Theor Appl Genet 123: 339–350

Allard R (1999) Principles of Plant Breeding, 2nd edn. John Wiley & Sons, New York

Bernardo R (1996) Best linear unbiased prediction of maize single-cross performance. Crop Sci 36: 50–56

Bernardo R (2010) Genomewide selection with minimal crossing in self-pollinated crops. Crop Sci 50: 624–627

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

Bordes J, Charmet G, de Vaulx RD, Lapierre A, Pollacsek M, Beckert M et al. (2007) Doubled-haploid versus single-seed descent and S_1-family variation for testcross performance in a maize population. Euphytica 154: 41–51

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, et al. (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet 6(5):e1000940

Buckler E, Holland J, Bradbury P, Acharya C, Brown P, Browne C, et al. (2009) The genetic architecture of maize flowering time. Science 325: 714–718

Choo T, Reinbergs E, Park S (1982) Comparison of frequency distributions of doubled haploid and single seed descent lines in barley. Theor Appl Genet 61: 215–218

Cornish M (1990) Selection during a selfing programme. I. The effects of a single round of selection. Heredity 65: 201–211

Cornish M (1990) Selection during a selfing programme. II. The effects of two or more rounds of selection. Heredity 65: 213–220

Courtois B (1993) Comparison of single seed descent and anther culture-derived lines of three single crosses of rice. Theor Appl Genet 85: 625–631

Daetwyler H, Pong-Wong R, Villanueva B, Woolliams J (2010) The impact of genetic architecture on genome-wide evaluation methods. Genetics 185: 1021–1031

Daetwyler H, Villanueva B, Woolliams J (2008) Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS One 3(10):e3395

Falconer D, Mackay T (1996) Introduction to quantitative genetics, 4th edn, Longman, Harlow, Essex, UK.

Farrar D, Glauber R (1967) Multicollinearity in regression analysis: the problem revisited. Rev Econ Stat 49: 92–107

Fess T, Kotcon J, Benedito V (2011) Crop breeding for low input agriculture: a sustainable response to feed a growing world population. Sustainability 3: 1742–1772

Goddard M, Hayes B (2007) Genomic selection. J Anim Breed Genet 124: 323–330

Guo Z, Tucker D, Lu J, Kishore V, Gay G (2012) Evaluation of genome-wide selection efficiency in maize nested association mapping populations. Theor Appl Genet 124: 261–275

Hallauer A, Miranda J (1988) Qunatitative genetics in maize breeding. Iowa State University Press, USA

Hastings W (1970) Monte carlo sampling methods using Markov chains and their applications. Biometrika 57: 97–109

Heffner E, Jannink J, Iwata H, Souzad E, Sorrells M (2011) Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51:2597–2606

Heffner E, Lorenz A, Jannink J, Sorrells M (2010) Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50:1681–1690

Heffner E, Sorrells M, Jannink J (2009) Genomic selection for crop improvement. Crop Sci 49:1–12

Hoeting J, Madigan D, Raftery A, Volinsky C (1999) Bayesian model averaging : a tutorial. Stat Sci 14:382–417

Iyamabo O, Hayes PM (1995) Effects of selection and opportunities for recombination in doubled-haploid populations of barley (hordeum vulgare l). Plant Breed. 114:131–136

Kearsey M, Sturley S (1984) A model for the incorporation of epistasis into a computer simulation for three experimental designs. Heredity 52: 373–382

Laurie C, Chasalow S, LeDeaux J, McCarroll R, Bush D, Hauge B, et al (2004) The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. Genetics 168:2141–2155

Lee S, van der Werf J, Hayes B, Goddard M, Visscher P (2008) Predicting unobserved phenotypes for complex traits from whole-genome SNP data. PLoS Genet 4(10):e1000231

Legarra A (2008) Performance of genomic selection in mice. Genetics 180: 611–618

Liu P, Zhu J, Lu Y (2004) Marker-assisted selection in segregating generations of self-fertilizing crops. Theor Appl Genet 109: 370–376

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

Luan T, Woolliams J, Lien S, Kent M, Svendsen M, Meuwissen T (2009) The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. Genetics 183: 1119–1126

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland

Mayor P, Bernardo R (2009) Genomewide selection and marker-assisted recurrent selection in doubled haploid versus F_2 populations. Crop Sci 49:1719–1725

McMullen M, Kresovich S, Villeda H, Bradbury P, Li H, Sun Q (2009) Genetic properties of the maize nested association mapping population. Science 325:737–740

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

Murigneux A, Baud S, Beckert M (1993) Molecular and morphological evaluation of doubled-haploid lines in maize. 2. comparison with single-seed-descent lines. Theor Appl Genet 87:278–287

Otto S, Jones C (2000) Detecting the undetected: estimating the total number of loci underlying a quantitative trait. Genetics 156:2093–2107

Park S, Walsh E, Reinbergs E, Song L, Kasha K (1976) Field perfomance of doubled haploid barley lines in comparison with lines developed by the pedigree and SSD methods. Can J Plant Sci 56:467–474

Piepho H (2009) Ridge regression and extensions for genomewide selection in maize. Crop Sci 49:1165–1176

Piyasatian N, Fernando R, Dekkers J (2007) Genomic selection for marker assisted improvement in line crosses. Theor Appl Genet 115(5):665–74

Pooni H, Jinks J (1985) Retrospective selection and sources of superior inbreds amongst pedigree inbred families of *Nicotiana rustica*. Heredity 54:255–260

Riggs T, Snape J (1977) Effects of linkage and interaction in a comparison of theoretical populations derived by diploidized haploid and single seed descent methods. Theor Appl Genet 49:111–115

Schaeffer L (2006) Strategy for applying genome-wide selection in dairy cattle. J Anim Breed Genet 123:218–223

Simmonds N (1979) Principles of crop improvement. Longman London, London

Snape J (1976) A theoretical comparison of diploidized haploid and single seed descent populations. Heredity 36:275–277

Verbyla K, Bowman P, Hayes B, Goddard M (2010) Sensitivity of genomic selection to using different prior distributions. BMC Proc 4(Suppl 1):S5.

Verbyla K, Hayes B, Bowman P, Goddard M (2009) Accuracy of genomic selection using stochastic search variable selection in Australian holstein friesian dairy cattle. Genet Res 91:301–311

Wells W, Weiser G (1989) Additive genetic variance within populations derived by single-seed descent and pod-bulk descent. Theor Appl Genet 78:365–368

Wong C, Bernardo R (2008) Genomewide selection in oil plam: increasing selection gain per unit time and cost with small populations. Theor Appl Genet 116:815–824

Zhong S, Dekkers J, Fernando R, Jannink J (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. Genetics 182:355–364