Received: 20 April 2023 | Accepted: 20 August 2023

DOI: 10.1002/tpg2.20388

The Plant Genome

ORIGINAL ARTICLE

Simulations of multiple breeding strategy scenarios in common bean for assessing genomic selection accuracy and model updating

Isabella Chiaravallotti[1] | **Jennifer Lin**[1] | **Vivi Arief**[2] | **Zulfi Jahufer**[2] | **Juan M. Osorno**[3] | **Phil McClean**[3] | **Diego Jarquin**[4] | **Valerio Hoyos-Villegas**[1]

[1]Department of Plant Science, McGill University, Montreal, Quebec, Canada

[2]School of Agriculture and Food Sustainability Faculty of Science, University of Queensland, Brisbane, Australia

[3]Department of Plant Sciences, North Dakota State University, Fargo, North Dakota, USA

[4]Agronomy Department, University of Florida, Gainesville, Florida, USA

Correspondence
Valerio Hoyos-Villegas, Department of Plant Science, McGill University, Montreal, QC, Canada.
Email: Valerio.hoyos-villegas@mcgill.ca

Assigned to Associate Editor Philomin Juliana.

Funding information
Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2020-07002; Fonds de recherche du Québec - Nature et technologies, Grant/Award Number: 322853

Abstract

The aim of this study was to evaluate the accuracy of the ridge regression best linear unbiased prediction model across different traits, parent population sizes, and breeding strategies when estimating breeding values in common bean (*Phaseolus vulgaris*). Genomic selection was implemented to make selections within a breeding cycle and compared across five different breeding strategies (single seed descent, mass selection, pedigree method, modified pedigree method, and bulk breeding) following 10 breeding cycles. The model was trained on a simulated population of recombinant inbreds genotyped for 1010 single nucleotide polymorphism markers including 38 known quantitative trait loci identified in the literature. These QTL included 11 for seed yield, eight for white mold disease incidence, and 19 for days to flowering. Simulation results revealed that realized accuracies fluctuate depending on the factors investigated: trait genetic architecture, breeding strategy, and the number of initial parents used to begin the first breeding cycle. Trait architecture and breeding strategy appeared to have a larger impact on accuracy than the initial number of parents. Generally, maximum accuracies (in terms of the correlation between true and estimated breeding value) were consistently achieved under a mass selection strategy, pedigree method, and single seed descent method depending on the simulation parameters being tested. This study also investigated model updating, which involves retraining the prediction model with a new set of genotypes and phenotypes that have a closer relation to the population being tested. While it has been repeatedly shown that model updating generally improves prediction accuracy, it benefited some breeding strategies more than others. For low heritability traits (e.g., yield), conventional phenotype-based selection methods showed consistent rates of genetic gain, but genetic gain under genomic selection reached a plateau after fewer cycles. This

Abbreviations: DF, days to flowering; GEBV, genomic estimated breeding value; GS, genomic selection; GWAS, genome-wide association study; LD, linkage disequilibrium; MAS, marker-assisted selection; PCA, principal component analysis; QTL, quantitative trait locus; rrBLUP, ridge regression best linear unbiased prediction; SNP, single nucleotide polymorphism; SY, seed yield; TBV, true breeding value; WM, white mold.

plateauing is likely a cause of faster fixation of alleles and a diminishing of genetic variance when selections are made based on estimated breeding value as opposed to phenotype.

1 | INTRODUCTION

1.1 | Genomic selection

First described by Bernardo (1994) to predict hybrid yields in maize based on restriction fragment length polymorphisms and later elaborated on by Meuwissen et al. (2001), genomic selection (GS) is a technique that uses information from genetic markers to predict the breeding value of selection candidates. Genomic prediction allows breeders to make selections based on the genomic estimated breeding values (GEBVs) of a pool of potential selection candidates. GEBVs are assigned using a prediction model trained with genotypes, and phenotypes from a training population and those with the most desirable GEBVs are advanced to the next generation. In order to accurately identify those individuals with a high GEBV, the prediction model must be properly informed by the training population. With advances in DNA technology and declining costs for genotyping, breeders can now gain access to high-density marker datasets. In particular, genome-wide association studies (GWAS) have allowed for the discovery of high-resolution genomic signals and quantitative trait loci (QTL) with significant effects for a given trait (Tibbs Cortes et al., 2021). Before the development of GS, genotypic data were first leveraged through marker-assisted selection (MAS). MAS is used to select individuals with favorable allele combinations by identifying markers that are closely linked to desired QTL. Much of the success from MAS has been in traits that are controlled by single genes of major effects. Application of MAS to polygenic traits, or traits controlled by many genes, has been less successful because the linkage phase between a marker and a QTL must be determined each time before implementing MAS (Assefa et al., 2019; Boopathi & Boopathi, 2020).

GS differs from MAS in that all markers across the genome are jointly used for prediction. It is assumed that with adequately high maker density, all causal loci will be linked to at least one marker (Voss-Fels et al., 2019). To implement GS, the first step is to build a training population, where individuals are both genotyped and phenotyped. This training population is used to train the prediction model. Then, the model is applied to a testing population, where individuals have only been genotyped, to obtain their GEBVs.

The advantage to using GS is that it has the potential to save the time and resources that would normally be dedicated to phenotyping individuals. It has also been proposed that new parents for the next breeding cycle can be selected in earlier generations, saving time and increasing genetic gain per unit time (Crossa et al., 2017). If gains from GS do not exceed those from "conventional" phenotype-based selection but are simply comparable, GS has the potential to reduce the financial load of a breeding program given the per-individual genotyping costs are lower than the per-individual costs for multi-environment yield trials (Bernardo, 2020). Further, depending on the genetic architecture of the trait, GS has potential to increase selection accuracy, with higher heritabilities generally resulting in greater accuracy (Massman et al., 2013). Genomic selection also allows breeders to identify poor performers as well as top performers. The ability to identify and discard poor performers can provide as much a benefit to breeders as identifying and keeping top performers. This saves the resources that would have been used to pinpoint poor performers in the field and quickly narrows the selection pool to only the most desirable candidates.

Genomic selection has been widely adopted in animal breeding programs. The breeding value of an individual is the sum of the average effects of the additive alleles it carries, which is determined by the performance of the individual's progeny (Nakaya & Isobe, 2012). To calculate the breeding value of a parent in animal breeding, it would first be necessary to produce and carefully phenotype several progenies over the course of several years.

GS has been prolific in animal breeding because it has allowed breeders to estimate breeding values quickly without first producing and evaluating progeny. Instead, selection candidates can simply be genotyped. This procedure drastically reduces the time and resources consumed in an animal breeding program, increasing gain per unit time. Adding to this effect is the fact that many traits of interest in animal breeding are complex (controlled by many small effect loci), so MAS provided limited benefit to animal breeders (Meuwissen et al., 2016). To provide an example, in dairy cattle, Doublet et al. (2019) found annual genetic gain increases of 33%–77% in total genetic merit index for three different breeds following the implementation of GS.

GS shows potential to increase selection accuracy, reduce cycle time, model complex quantitative traits, and model genotype by environment interactions in plants (Crossa et al., 2017). However, the move toward implementing GS in plants, especially for complex quantitative traits, has been slow. Only a handful of studies have investigated the empirical benefit of GS in crop breeding programs (Beyene et al., 2015; Dreisigacker et al., 2023; Das et al., 2020; Massman et al., 2013; Zhang et al., 2017), and very few have directly

compared GS to conventional phenotypic selection (Bandillo et al., 2023). Contributing to this slower adoption of GS in plants is the challenge of constructing a model that can accurately estimate GEBVs from genotypic data. There are several factors that affect the accuracy of GS models, some of which are discussed below.

Core Ideas

- Trait architecture and breeding strategy had larger impact on genomic selection accuracy than initial number of parents.
- Some breeding strategies benefitted from model updating more than others.
- For yield genetic gain, phenotype selection showed consistent rates, but genomic selection plateaued faster.

1.2 | Factors impacting GS accuracy

1.2.1 | Training population, trait heritability, and linkage disequilibrium

A larger training population has been shown to generally increase prediction accuracy. An adequately large training population is especially important when the relatedness between the training and testing population decreases (Edwards et al., 2019; Zhang et al., 2017). The ability of the training dataset to effectively inform a model depends on several factors including marker density, population size, and relatedness to the population that the model will ultimately be applied to. In general, as relatedness between the training and testing population decreases, marker density must increase. Additionally, larger training populations tend to result in more accurate estimations of GEBVs (Jannink et al., 2010). Linkage disequilibrium can also impact model performance. Two highly related populations, which share large haplotype blocks and have high linkage disequilibrium (LD), may function well as a training/testing pair, but when the model is applied to another population with different LD patterns, performance will diminish (Verges & Van Sanford, 2020).

Furthermore, the heritability of a trait can impact the training population size required for accurate predictions, especially when the h^2 is less than 0.4. For example, to obtain an accuracy of 0.7, a training population size of 9000 is required for a trait with $h^2 = 0.2$ if the effective population size, or size of an ideal population required to adequately represent genetic variance in the population under consideration, is 1000. This greatly contrasts a training population size of 3000 when the trait heritability is 0.5 (Lorenz et al., 2011). While conventional breeding programs already phenotype large populations in order to make selections, the limiting factor in formulating a training population will likely be the cost of genotyping.

1.2.2 | Model update

Marker effects must be re-estimated during a GS breeding program to maintain accuracy and maximize gains. Accuracy is greater when the training population contains individuals in the same generation as the selection candidates (Heffner et al., 2010). A simulation study based on a sorghum breeding program found that updating the GS model every year can increase genetic gains up to 39% (Muleta et al., 2019). In a GS study simulating 60 years of selection on conifer trees, it was found that without updating the model, GS outperformed phenotype-based selection in the short term, but not in the long term. However, when a model update was implemented, GS gains were consistently higher than gains from phenotype-based selection (Iwata, 2011). In essence, as the number of generations separating the relatedness between training population and selection candidates increases, the accuracy will decrease, so updating the model is essential. A program's ability to regularly update the model will largely depend on the resources available for both regular genotyping and phenotyping. Exemplified in Rutkoski et al. (2015b), when initiating a GS pipeline, breeding programs will typically employ historic data to train the initial model, and then use data from the new breeding cycle to generate training data for subsequent model training and predictions. Model updating in this example led to a 1.86× increase in prediction accuracy, where a decrease of 0.02–0.31× in prediction accuracy would have been expected. After the model is updated, it is recommended to discard historical data (Rutkoski et al., 2015a); however, the breeder will ideally optimize the training set to suit their particular testing population (Berro et al., 2019; Isidro et al., 2015).

1.2.3 | Marker density

Genomic selection assumes that all causal QTLs are in linkage disequilibrium with at least one of the single nucleotide polymorphism (SNP) markers used for genomic prediction. Therefore, the density of SNP markers has an impact on prediction accuracy. There is no one-size-fits all for the number of SNPs required to achieve high prediction accuracy. The ideal number of SNPs depends on linkage decay in the population and the number of segregating segments in the genome. A minimum of N_eL SNPs (N_e = effective population size, L = number of segregating segments) are required to achieve moderate accuracy (Meuwissen, 2009). Accuracy tends to

plateau after a certain marker density is reached. For example, Haile et al. (2018) investigated the impact of marker density on genomic prediction accuracy in wheat. They tested a subset of 100, 500, 2000, 3000, 4000, and 9000 SNPs sampled from a total of 9752 SNPs. Three thousand SNPs were required for peak accuracy in predicting protein content, 2000 were required for gluten index, and 4000 were needed for grain yield. Asoro et al. (2011) investigated the impact of marker density on genomic prediction in wheat by sampling 300, 600, and 900 markers for model training. They also found that the ideal marker density varies by trait with some traits plateauing at 600 markers and others continuing to increase up to 9000 markers.

1.3 | Effectiveness of GS in plant breeding

Although GS has been widely implemented in animal breeding, its use in plant breeding still requires empirical validation. The handful of empirical GS studies carried out in plants have found an advantage to using GS over MAS in maize (Massman et al., 2013), an advantage to using GS over traditional phenotype-based selection in soybean and maize (Bandillo et al., 2023; Beyene et al., 2015), and a higher performance of GS lines over hybrid checks under abiotic stress in maize (Das et al., 2020). Genomic prediction also shows promise in dry beans in general (Keller et al., 2020; Shao et al., 2022), but evidence is limited. If GS is going to be successfully implemented in a dry bean breeding program as a technique for increasing the rate of genetic gain, simulation studies will be an important first step. Lin et al. (2022) carried out series of simulations to test five selection strategies (bulk breeding, single seed descent, mass selection, the pedigree method, and the modified pedigree method), three breeding frameworks (conventional, GS, and speed breeding), four parental population sizes (15, 30, 60 and 100), and three traits (seed yield [SY], days to flowering [DF], and white mold [WM] tolerance) in common bean. Expanding on the previous common bean simulation study, the objective of this study was to investigate the accuracy of GS in a simulation study in dry beans. As in Lin et al. (2022), five breeding strategies were simulated with the selection on three traits. The following hypotheses were tested:

1. Selection strategy influences GS accuracy and genetic gain.
2. GS model updating influences genetic gain and prediction accuracy to varying degrees depending on selection strategy, trait heritability, and number of initial parents in the breeding program.
3. By accounting for population size, heritability, and number of independent loci, GS accuracy can be predicted prior to implementation.

2 | METHODS

2.1 | Simulation parameters

In this experiment, five breeding strategies (mass selection, bulk breeding, single seed descent, pedigree method, and modified pedigree method) with a parental population size of 30 were simulated using the QuLinePlus module (Hoyos-Villegas et al., 2019). QuLinePlus is a module in the stochastic simulation platform Qu-Gene, which allows breeders to simulate different breeding program parameters. The population was developed following a consensus map reported by Galeano et al. (2011). This map used recombinant inbred lines from three different mapping populations in the Mesoamerican gene pool. Because a consensus map allowed for the identification of more markers than would a single population, this map was ideal for conducting these simulations. The map has 1010 markers and a length of 2041 cM. Selection was carried out on three traits: DF, WM tolerance), and SY. Because they represent the heritable portion of the genome which is accounted for when a breeding value is estimated, only additive effects were assumed for the QTL associated with each trait. Because QuLinePlus simulates populations in Hardy–Weinburg equilibrium with negligible or no LD, LD was generated using the forward-in-time population simulation software SimuPOP. In brief, the population underwent natural selection in SimuPOP, generating adequate linkage blocks. Then, the population files were converted to QuLinePlus-compatible files to conduct the simulations. An LD analysis can be viewed in Lin et al. (2022).

The simulation was conducted with 20 runs, each with six cycles, each cycle consisting of eight generations. Breeders aim to increase SY, but yield components are generally challenging to breed due to their complex quantitative nature, and the fact that these components may vary genetically between different gene pools or market classes of common bean. WM poses a major threat to common bean crops, especially in Europe, and North America and breeders typically focus on accumulating specific WM-resistant genes. When it comes to DF, breeders generally aim to stabilize flowering time in the face of abiotic stresses (Assefa et al., 2019).

The five breeding strategies will be explained here in brief. For further detail, see fig. 1 in Lin et al. (2022). In mass selection, top candidates are identified from an entire population (as opposed to a single family) and, for ease of implementation, advanced to the next generation without maintaining family structure. These selections are repeated for many generations until the multi-environment trial phase. In bulk breeding, a population resulting from crossing two parents is bulked and advanced until the population has reached the desired level of homozygosity. Then, selections are carried out, usually starting at the F_6 generation when yield trials are conducted. In single seed descent, single seeds are advanced

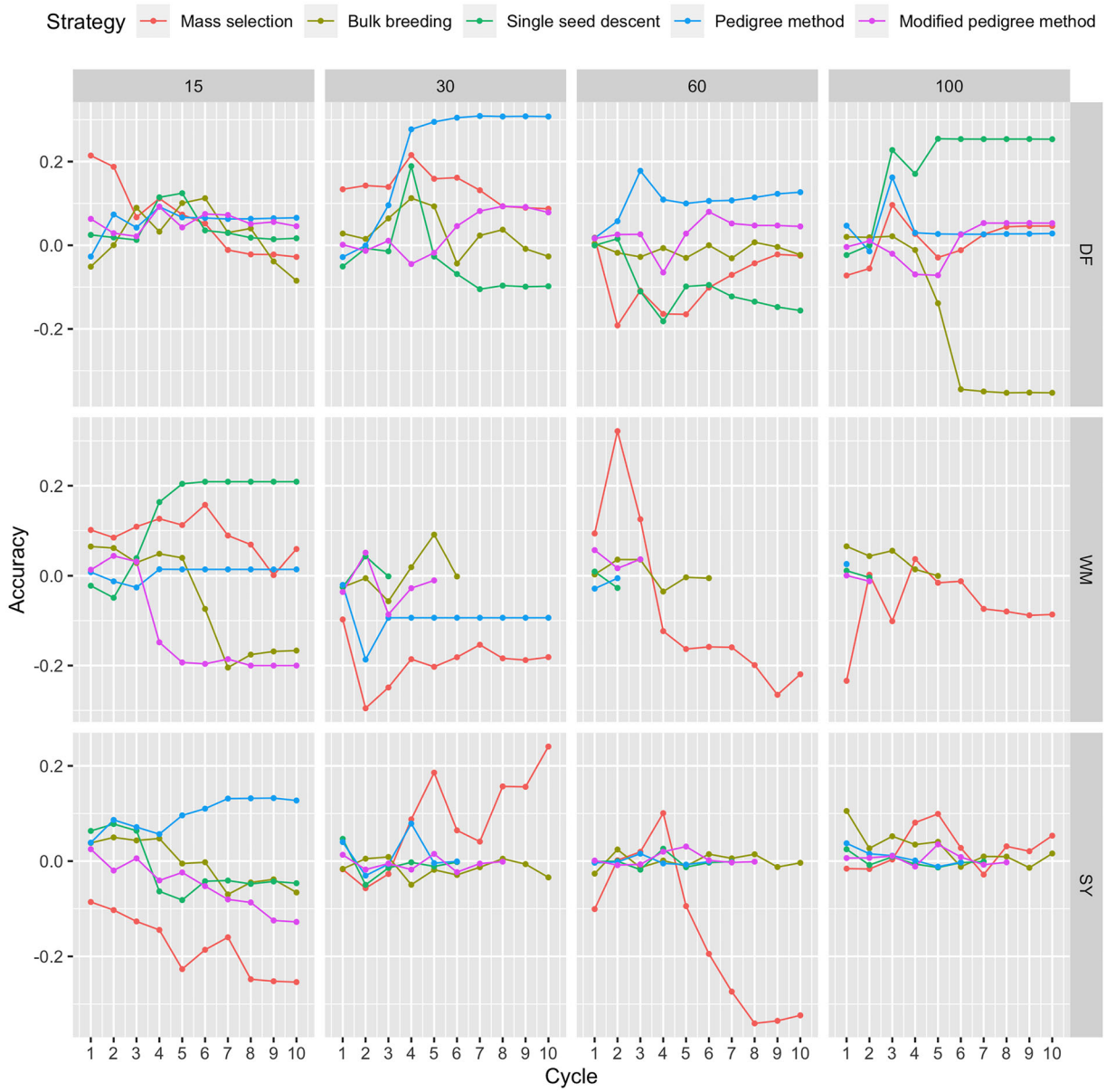

FIGURE 1 Expected estimated genomic selection accuracies calculated from Equation (3). Colored lines correspond to breeding strategies, which include mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method. Three traits were selected with differing parental population sizes indicated at the top and right-hand side of the panels. Traits included days to flowering (DF), white mold tolerance (WM), seed yield (SY).

to the F_5 generation at which point top candidates are selected and trialed. In the pedigree method, selection begins at the F_2, where lines are derived from single plants, forming $F_{2:3}$ families. Then, the top rows are selected from $F_{2:3}$ families to form $F_{3:4}$ families. Plants are then grown in family rows, and entire family rows are selected to advance to the next generation. For the modified pedigree method, individual plant selections take place in the $F_{2:3}$ and $F_{3:4}$ generations where winter nurseries are used in the $F_{3:4}$ and $F_{3:5}$ generation. Seeds are bulked in the winter nurseries, shortening the breeding cycle. Field trials start with $F_{3:6}$ seed.

2.2 | Implementing the GS model

The initial parental population was used as the initial training population. As described above, we created an initial parent population using SimuPOP and converted the population data to a format compatible with QuLinePlus. Using the simulated genotypes and phenotypes generated in QuLinePlus, the initial model was trained, and marker effects were obtained. GEBVs were used to make early-cycle selections (F2, F3, and F4 generations) following each respective selection strategy described in the previous section. In later generations,

when field trials were simulated, we made selections based on phenotype. In other words, early-cycle selections were made using GEBVs following each respective selection strategy. Late-cycle selections and parent selections for the next cycle were made using phenotype (SY) during simulated yield trials.

The ridge regression best linear unbiased prediction (rrBLUP) model was selected for estimating GEBVs (Gianola, 2013). This model has become the standard for GS because datasets used for model training and prediction tend to have far fewer observations (individuals) than parameters (markers), and rrBLUP accounts for this overparameterization. Because we have simulated additive effects, a linear model was the most appropriate choice. To implement the model, the R package "rrBLUP" (Endelman, 2011) was used to determine the marker effects for estimating GEBVs. The linear predictor (model) implemented in rrBLUP is shown in Equation (1):

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{y}$ is a vector of phenotypes, $\mathbf{X}$ is a design matrix for the fixed effects $\boldsymbol{\beta}$, $\mathbf{Z}$ is a design matrix for the random effects $\mathbf{u}$, or marker effects, where $u \sim N(0, I\sigma^2{}_u)$, and ε is residual variance. The genotypes and phenotypes were obtained from the parental population generated using the simuPOP program (Peng & Kimmel, 2005). Genomic selection was simulated by using the mixed.solve function in the R package rrBLUP to estimate the effect of each marker on the phenotype. Once the marker effects were determined, they were input into the "MarkerEffects" sheet with a genotyping error rate of 5%. From there, individuals were selected based on the sum of the linear combination between markers and their corresponding marker effects (genomic estimated breeding value, as opposed to conventional breeding in which selections are made on phenotype).

2.3 | GS model updating

In a previous study, Lin et al. (2022) obtained GS accuracies from simulations among various selection strategies. However, the model reported in Lin et al. (2022) was not updated during the simulation. To simulate updating the GS model, a parental population was used as the reference population in the first three cycles by using SimuPOP (Peng & Kimmel, 2005). At the end of the third cycle, a sample of selected individuals from the third cycle were used as the reference population to estimate new marker effects. Previous training data were discarded. From cycle four to six, selection was conducted on the basis of the newly estimated marker effects. GS with and without model updating was simulated under each of the scenarios described in the Section 2.1 (five selection strategies and three traits with 20 runs and six cycles).

2.4 | GS model accuracies

GS accuracies were assessed via two estimates: first, based on Equation (2) (expected accuracy) (Daetwyler et al., 2010); second, based on the correlation between the true breeding value (TBV) and GEBV (in silico realized accuracy). Using these two measures of accuracy, the influence of selection strategy and genetic architecture on GS model performance was assessed.

2.4.1 | Expected accuracy

Expected accuracy was calculated using Equation (2):

$$r_{g\hat{g}G} = \sqrt{\frac{N_p h^2}{N_p h^2 + M_e}}. \quad (2)$$

Daetwyler et al. (2010) described several components that impact the accuracy of GS. The authors derived a formula for GS accuracy, as follows:

$$r_{g\hat{g}G} = \sqrt{\frac{N_p h^2}{N_p h^2 + n_G}}, \quad (3)$$

where N_P refers to the number of individuals in a training population, h^2 is the narrow sense heritability, and n_G is the number of independent loci. Based on the derived formula, the accuracy of GS is influenced by the heritability of the trait, the number of individuals in the training population, and the number of loci being considered. As LD will result in some of the loci being linked, the number of independent chromosome segments (M_e) should be used in place of n_G (Daetwyler et al., 2010). The parental population generated using simuPOP will have adequate LD, which must be accounted for. By replacing n_G with M_e, one can derive Equation (2):

$$r_{g\hat{g}G} = \sqrt{\frac{N_p h^2}{N_p h^2 + M_e}}, \quad (2)$$

where N_P refers to the number of individuals in a training population, h^2 is the heritability, and M_e is the number of independent chromosome segments. Equation (2) was the formula used to obtain predicted GS accuracies. Equation (4) was used to calculate M_e:

$$M_e = \frac{2N_e L}{\log(4N_e L)}, \quad (4)$$

where N_e is the effective population size and L is the genome length in Morgans. Effective population size can be defined as the size of an ideal population undergoing the same rate of drift as the population at hand (Waples, 2022).

Equation (4) requires the estimation of N_e. To accomplish this, the variance effective size estimator described by (Waples, 1989) was used and is shown in Equation (5):

$$\hat{N}_e = \frac{t}{2\left[\hat{F}_c - \left(\frac{1}{2S_0} + \frac{1}{2S_t}\right)\right]}, \quad (5)$$

where t refers to the number of generations that have elapsed between the two sampled populations, $\hat{F}_c$ refers to the estimator for the standardized variance of gene frequency changes at a single locus, S_0 and S_t indicate the sample sizes of the population at time 0 and time t, respectively. Because a breeding population is not isolated enough to reach an equilibrium between drift and mutation, we are able to calculate this "local" effective population size (as opposed to a more generalized "global" effective population size) using the two sample populations in Equation (5) to serve as an experimental basis for our GS accuracy estimations (Waples, 2022).

The estimator Fc can be written as:

$$F_c = \frac{1}{k}\sum_{i=1}^{k}\frac{\left(x_i - y_i\right)^2}{\left(x_i + y_i\right)/2 - x_i y_i}, \quad (6)$$

where k is the number of alleles, x_i is the observed allele frequency at time 0, and y_i is the observed allele frequency at time t.

In our study, average F_c estimates for all loci were determined and used to determine N_e. From there, GS accuracy was estimated using Equation (3). One of the output files generated by QU-GENE is the *fre* file, which provides the allele frequencies at each marker. These allele frequencies were entered into Equation (6) and subsequently used to calculate the N_e in Equation (5), M_e in Equation (4), and finally the expected accuracy in Equation (2). All calculations were performed using original code written in R, which can be found on the McGill University Pulse Breeding and Genetics Laboratory GitHub page (see Data Availability Statement).

2.4.2 | In silico realized GS accuracy: TBVs and GEBVs

Simulation outputs were used to estimate the accuracy of GS. More specifically, the breeding values were obtained from the *.pou* files generated by the QuLinePlus Module. The genotypic values from the conventional phenotype-based breeding framework modeled in QuLinePlus were used as TBVs. The genotypic values from the GS framework represented the GEBVs.

QU-GENE performs selection on the per-plot basis. Thus, the correlation between the family mean GEBV and TBV values was used as a direct estimate of realized GS accuracy. Original code was used to conduct calculations and correlations and may be found on the lab GitHub (McGill Pulse Breeding and Genetics Laboratory, 2022).

2.5 | Principal component analysis

Principal component analyses were conducted to visualize the relationships between the factors that might influence both genetic gain and GS accuracy. The family means for seven different factors that contribute to the genetic gain and GS accuracy in the first cycle were determined for each of the 20 runs. The seven factors included genetic gain, fixation of favorable alleles, Hamming distance, genetic variance, effective population size, TBV, and genomic estimated breeding value. The fixation of favorable alleles described the average percentage of beneficial alleles that were fixed in the population. Meanwhile, the Hamming distance was used to describe the distance of an individual from an ideal genotype. This distance was determined as the number of base pairs that differ from the optimal genotype. The effective population size was calculated according to Equation (5). All calculations were performed using original code written in R and may be located on the lab GitHub page (see Data Availability Statement). Lastly, the R packages ggbiplot (version 0.55) and ggplot2 (version 3.3.3) were used to create the principal component analyses (PCA).

3 | RESULTS

3.1 | Genomic estimated breeding values

GEBVs were determined for each cycle for 10 cycles (Figure S1). Over the course of 10 cycles, GEBVs saw either a gradual increase or decrease before reaching a plateau. For DF and SY, the GEBVs increased rapidly before plateauing. The opposite trend was observed for WM tolerance, with GEBVs declining before leveling off. The parental population sizes had an impact on the GEBVs at the end of the breeding program. For DF, the GEBVs averaged across the strategies were 19.90, 17.47, 13.98, and 9.87 for 15, 30, 60, and 100 parents, respectively. For WM tolerance, parental population sizes of 15, 30, 60, and 100 resulted in average GEBVs of −35.26, −61.52, −62.34, and −62.52, respectively. Lastly, for SY, the average GEBVs were 2420.18, 2745.80, 2185.04, and 2183.38 for parental population sizes of 15, 30, 60, and 100, respectively.

3.2 | True breeding values

TBVs were obtained from the QU-GENE output files and plotted over 10 cycles (Figure 2). Over the course of 10 cycles, TBVs saw either a gradual increase or decrease before reaching a plateau. For DF and SY, the TBVs increased and eventually plateaued. The opposite was true for WM tolerance, where TBVs declined before reaching a plateau. There were notable differences between the TBVs when different numbers of parents were used at the beginning of the cycle. For each of the traits, as the number of parents increased, the average TBV for the strategies decreased. For DF, the average TBVs across strategies at the end of the 10th cycle were 20.21, 17.69, 13.95, 9.96 DF for 15, 30, 60, and 100 parents. For WM (0–100 severity score) tolerance after 10 cycles, the average TBVs were −46.17, −62.32, −62.51, and −62.55 for 15, 30, 60, and 100 parents, respectively. For SY, the average TBVs were 2830.95, 2759.40, 2190.16, and 2189.79 kg/ha for 15, 30, 60, and 100 parents, respectively. For most breeding scenarios, bulk breeding, single seed descent, pedigree, and modified pedigree methods led to similar TBVs. Mass selection resulted in a lower TBV for DF and SY, while it led to a higher TBV for WM tolerance in comparison to the other four strategies.

3.3 | GS model accuracies

3.3.1 | Expected GS accuracy

GS accuracies determined using Equation (3) ranged from 0.07 for SY to 0.63 for DF (Figure 1). In general, prediction accuracy decreased over the 10 cycles. The decline was smaller with parental population sizes of 15 and 30. Prediction accuracies were higher with larger parental population sizes. The strategies had similar accuracies and followed similar trends when the parental population size was small. However, with large parental population sizes, mass selection had a much greater prediction accuracy compared to the other strategies. Furthermore, the accuracy remained relatively high for mass selection. The accuracy was highest under DF, followed by WM tolerance and then SY. For DF under mass selection with 100 parents, the accuracy decreased from 0.63 to 0.47 over 10 cycles. Meanwhile, for WM tolerance under mass selection with 100 parents, the accuracy decreased from 0.46 to 0.39 over 10 cycles. Lastly, for SY under mass selection with 100 parents, the accuracy decreased from 0.43 to 0.29 over 10 cycles. In most breeding scenarios, bulk breeding resulted in the lowest prediction accuracies. For DF with 15 parents, the accuracy in bulk breeding decreased from 0.18 to 0.10 over 10 cycles. For WM tolerance with 15 parents, accuracy declined from 0.11 to 0.09 over 10 cycles when bulk breeding was used. For the selection of SY with 15 initial parents, accuracy with bulk breeding decreased from 0.09 to 0.07 over 10 cycles. Heritability had an impact on GS accuracy, where accuracy was the highest under DF, followed by WM tolerance and then SY. However, selection strategies had similar accuracies when the parental population size was small, regardless of heritability.

3.4 | In silico realized GS accuracy

GS accuracies were estimated from the correlation between the TBV and the GEBV. Under each of the five strategies, accuracies saw a general decline before plateauing. They ranged from −0.35 for SY to 0.32 for WM (Figure 2). The mean accuracies for each strategy were −0.03, −0.02, 0.02, 0.05, and −0.01, for mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method, respectively.

For DF, the highest accuracy (0.31) was observed under the pedigree method with 30 parents, while the lowest accuracy (−0.35) was in bulk breeding with 100 parents. When considering WM, single seed descent with 15 parents resulted in the greatest accuracy (0.32). The lowest WM accuracy was seen under mass selection with 60 parents (−0.29). Interestingly, in cycle 2, there was an increase in accuracy, after which the WM accuracy declined rapidly and became negative by cycle 4. For SY, both the highest (0.24) and lowest (−0.34) accuracies were observed in mass selection. For certain cycles, a correlation could not be obtained. In these cycles, the variance was zero, and the correlation was undefined.

3.4.1 | GEBV with model update

GEBVs were plotted over six cycles (Figure 3). For DF and SY, GEBVs showed an increase following the model update. For WM, GEBVs showed an increase and a subsequent decline. For DF, there was an increase from cycles 3 to 4 for mass selection, the pedigree method, and the modified pedigree method, with increases of 19.0, 18.8, and 20.6. Smaller increases were observed for the other two strategies. GEBVs increased by 7.66 and 6.02 between cycles 3 and 4 for bulk breeding and single seed descent, respectively. For WM tolerance, a large increase was observed for bulk breeding, while single seed descent led to the smallest increase in GEBV. From cycles 3 to 4 for WM tolerance, GEBVs increased by 4.03, 39.1, 9.43, 16.7, and 13.1 for mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method, respectively. Lastly, for SY, all five strategies resulted in an increase in GEBV following model update, with the greatest increase observed in mass selection and the smallest increase in single seed descent.

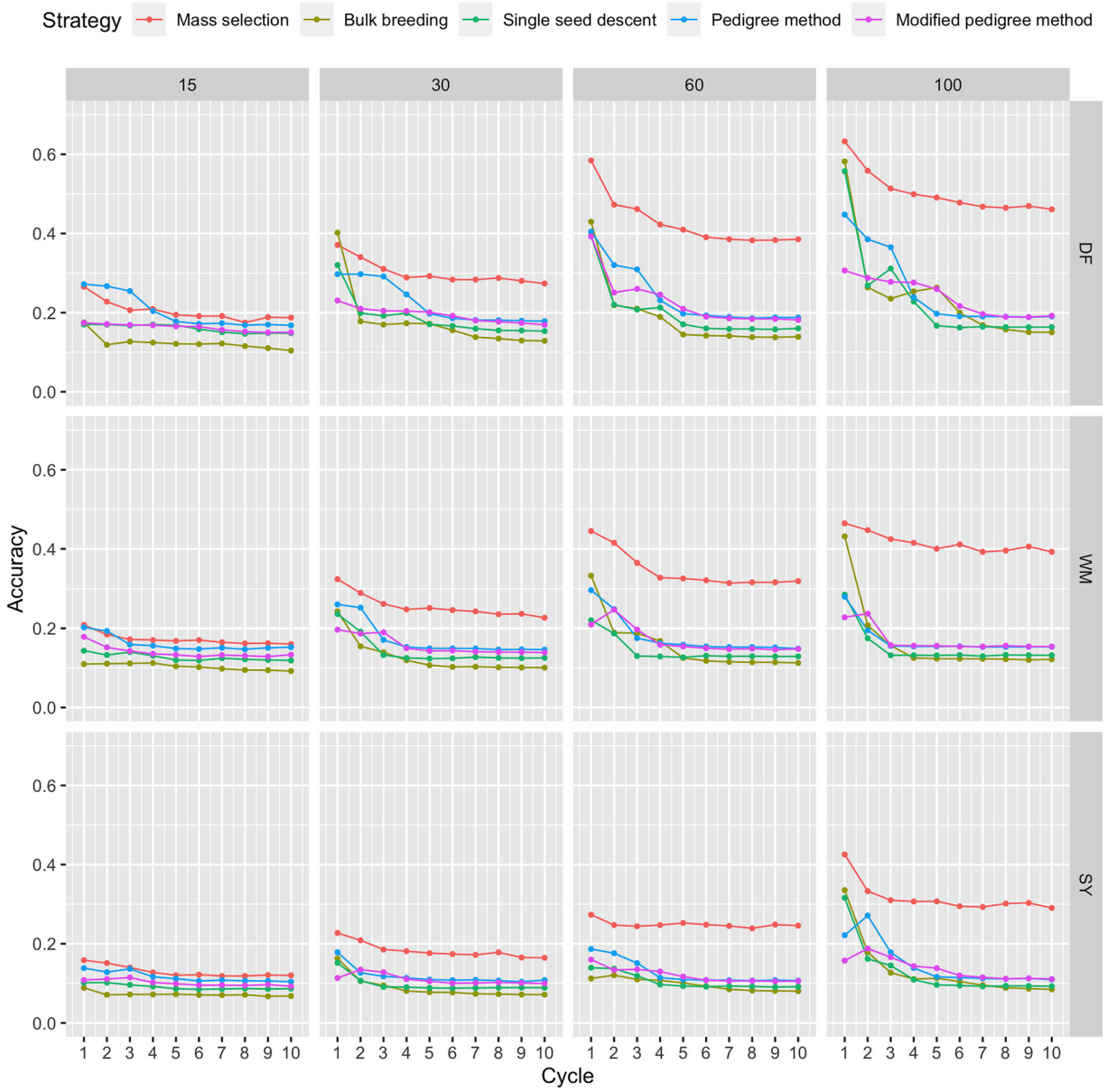

FIGURE 2 In silico realized genomic selection accuracy estimated from QU-GENE output for three selected traits: days to flowering (DF), white mold tolerance (WM), and seed yield (SY). Accuracies were calculated as the correlation between the true breeding value and the genomic estimated breeding value.

GEBVs increased by 1867, 367, 130, 886, and 1251 for mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method, respectively.

3.4.2 | TBVs with model update

After model update, the TBVs were determined and plotted over six cycles (Figure 3). In general, updating the model resulted in an increase in TBVs. At cycle 3 (where the update occurred), there was an increase in the TBV for all breeding scenarios. For DF, the TBV increased by 9.53, 7.05, 6.32, 4.34, and 6.38 from cycles 3 to 4 for mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method, respectively. For WM tolerances, TBVs rose by 15.0, 38.3, 9.44, 0.73, and 10.7 from cycles 3 to 4 for mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method, respectively. Lastly, for SY from cycles 3 to 4, mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method had increases in TBVs of 761, 299, 134, 129, and 134, respectively. TBVs appeared to plateau after cycle 4 for DF and SY. However, for WM tolerance, TBVs rapidly declined after cycle 4.

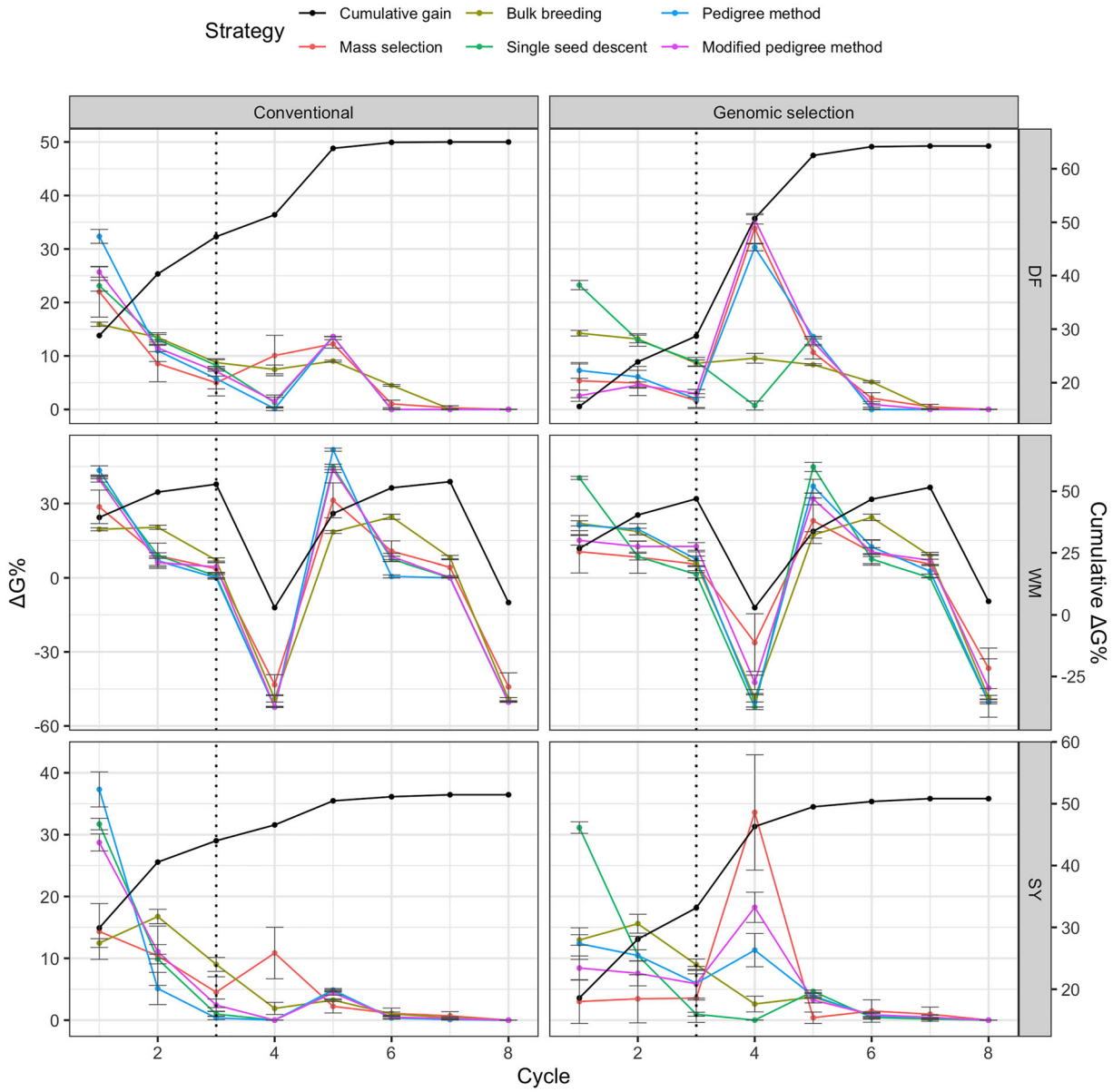

FIGURE 3 True breeding values (TBV) and genomic estimated breeding values (GEBV) plotted over six cycles for three traits in simulation with genomic selection (GS) model update. Traits include days to flowering (DF), white mold tolerance (WM), and seed yield (SY). Vertical dotted line indicates the point at which the GS model was updated.

3.4.3 | Expected GS accuracy with model update

GS accuracies determined using Equation (3) ranged from 0.08 for SY to 0.58 for DF (Figure 4). For all breeding scenarios, a general trend was observed where an increase in accuracy occurred after the GS model update at cycle 3, followed by a decline from cycles 4 to 5. The peak accuracy predicted from the DF simulation was 0.58, occurring at cycle 4 with mass selection. For WM tolerance, the peak accuracy was 0.51, occurring at cycle 4 with the pedigree method. The peak accuracy for SY was 0.38 at cycle 4 using the pedigree method.

3.4.4 | In silico realized GS accuracy with model update

In addition, the in silico realized accuracies fluctuated from one cycle to the next regardless of the model update. GS accuracies, represented by the correlations between the TBV and the GEBV, were obtained and plotted over six cycles (Figure 4). Updating the model generally did not improve accuracies. Once again, the accuracy fluctuated over the different cycles. Mass selection had the greatest variability in accuracy, in some cycles having the highest accuracy, while in others having the lowest accuracies. For DF, following the model update at cycle 4, there was a small improvement

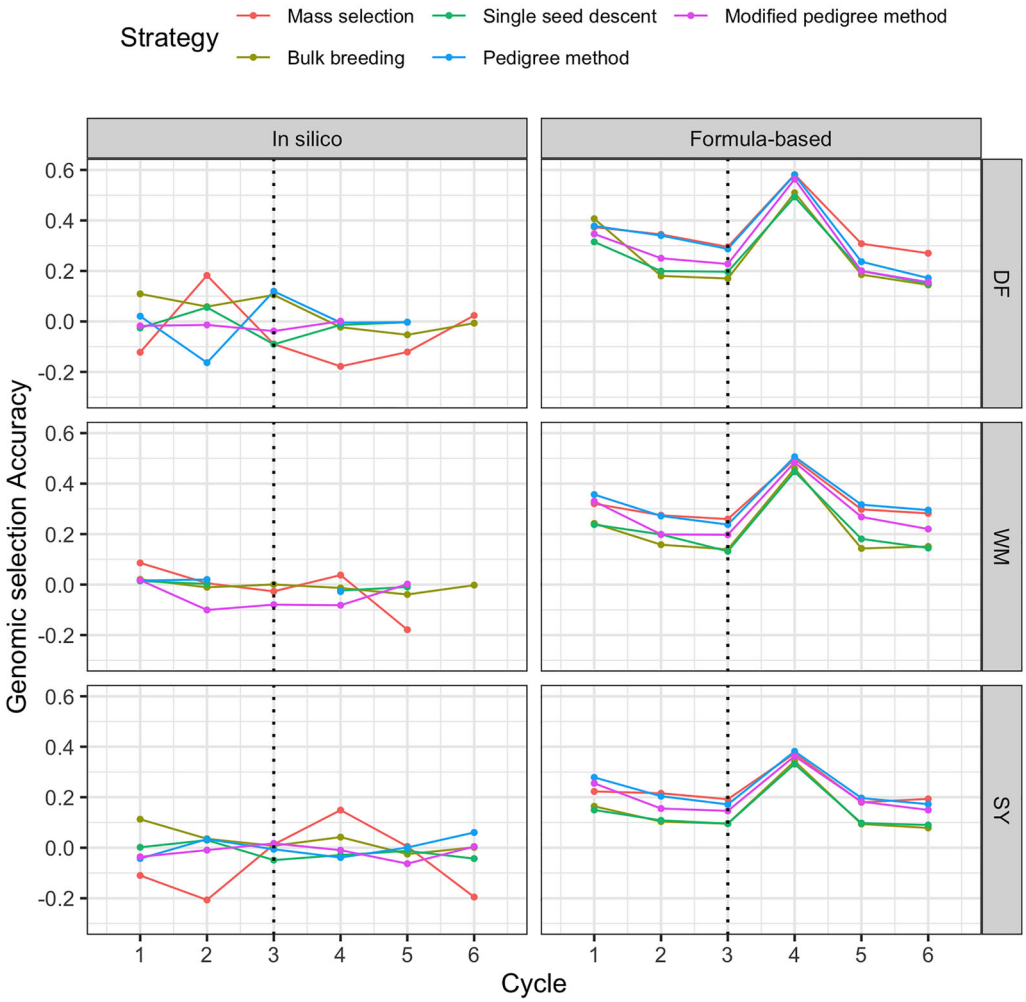

FIGURE 4 In silico realized and expected estimates of genomic selection accuracy after model update. Vertical line indicates cycle in which the GS model was updated. Prediction accuracies are shown over six cycles. Colored lines correspond to strategies, which include mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method. In silico estimates correspond to the correlation between BV and GEBV at each cycle. Formula-based estimates were obtained using Equation (3).

in accuracy under single seed descent and the modified pedigree method, where accuracies increased by 0.08 and 0.04, respectively. The other three strategies experienced a decrease in accuracy when predicting genetic values for DF.

From cycles 3 to 4, WM tolerance GS accuracies declined by 0.09, 0.13, and 0.12 for mass selection, bulk breeding, and the pedigree method, respectively. GS accuracies increased by 0.14, 0.03, and 0.02 between cycles 3 and 4 for mass selection, bulk breeding, and single seed descent, respectively. Although mass selection resulted in an increase in accuracy after cycle 3, it rapidly dropped and became negative. Decreases in GS accuracy after the third cycle were observed for the pedigree method and the modified pedigree method. However, the strategy with the highest accuracy in the final cycle was the pedigree method, with a value of 0.06.

3.4.5 | Genetic gain with model update

The results from the model update (Figure 5) indicated that there was a sharp increase followed by a rapid decline in genetic gain. Model update only improved genetic gain in one or two cycles immediately after the update, only to return to the rates of genetic gain prior to the update. Conventional breeding was included alongside GS as a comparison for model update. Figure 5 shows that updating the GS model resulted in an increase in genetic gain after cycle 3 for mass selection, the pedigree method, and the modified pedigree method when selecting for DF and SY. However, it led to a decrease in genetic gain immediately after cycle 3, followed by an increase after cycle 4, and a decrease after cycle 5 for all strategies when selecting for WM tolerance. When compared to conventional phenotype-based breeding, GS led to higher levels of genetic gain for certain strategies in the cycle after

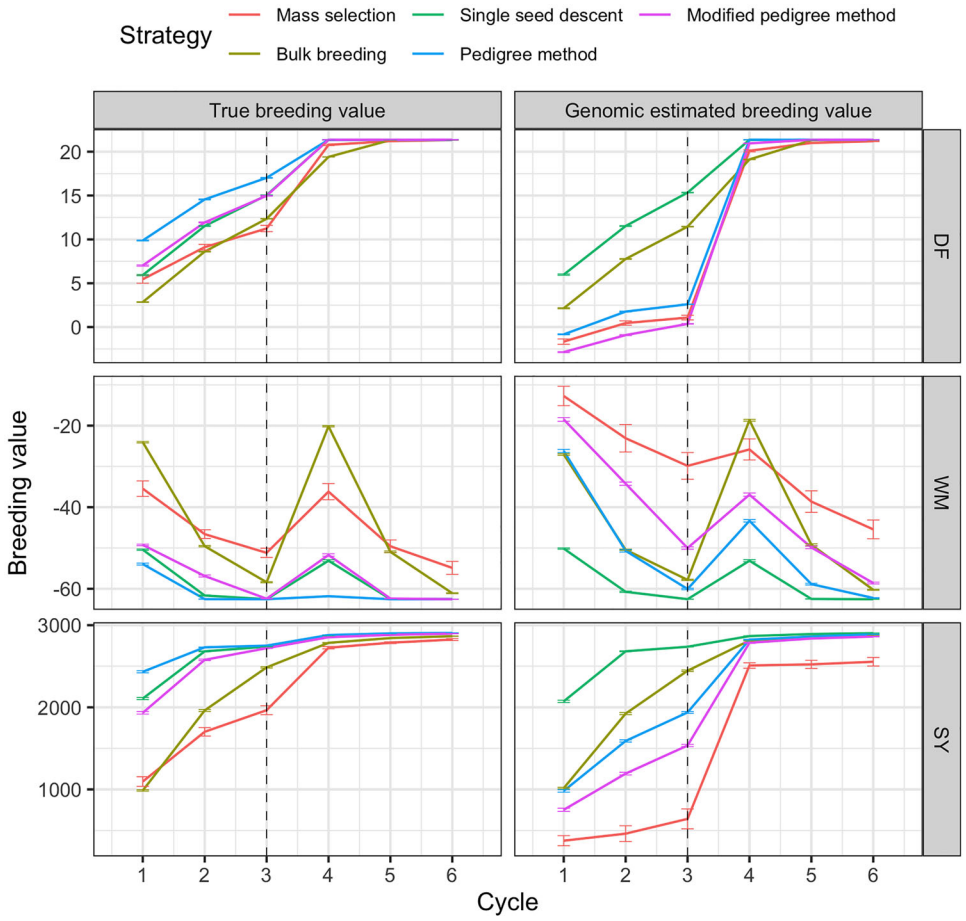

FIGURE 5 Comparison of five breeding strategies in terms of relative genetic gain following model update. Conventional breeding was included as a control for interruption of the simulation run. Colored lines correspond to the breeding strategies, mass selection, bulk breeding, single seed descent, pedigree method, and modified pedigree method, with the appropriate scale shown on the left of the plot. Black line indicates the cumulative genetic gain averaged across the five strategies, with the corresponding scale to the right of the plot. Simulated traits included days to flowering (DF), white mold tolerance (WM), and seed yield (SY). Dotted line shows when model update took place.

the GS model was updated. For DF, mass selection under GS was 23.8% higher compared to conventional phenotype-based breeding. Meanwhile, the pedigree method and the modified pedigree method were 30.2% and 34.0% higher in GS, respectively. For WM tolerance, mass selection led to 17.0% greater genetic gain using GS than conventional breeding, while the modified pedigree method using GS resulted in 9.94% higher genetic gain. Finally, for SY, mass selection, the pedigree method, and the modified pedigree method resulted in 22.7%, 11.3%, and 18.2% higher genetic gain, respectively, using GS compared to conventional breeding. For all other breeding scenarios, there was little to no difference between GS and conventional breeding in the cycle after the GS model update.

3.5 | Principal component analysis

Principal component analysis for each trait and breeding strategy was conducted to show the overall result of the simulation with the model update. For each trait, a PCA was performed on a matrix of with each breeding strategy, simulation run, and breeding cycle with values for genetic gain, fixation of favorable alleles, hamming distance, effective population size, TBV, and GEBV. The analysis was colored according to the breeding strategy so that the impact of the breeding strategy could be observed. Figure 6 shows the PCA plot for DF, where 80.53% of the variance is explained by the first two principal components. Notably, the eigenvectors for TBV and GEBV are close together and point in the same direction. The eigenvector for genetic gain and Hamming distance point in similar directions. Toward the right side of the PCA plot, there were two clusters for mass selection that formed on the extreme of the GEBV and TBV eigenvectors. On the opposite side, a cluster containing all five strategies was found in the extremes of both the Hamming distance vector and the genetic gain vector. In the direction of the eigenvector for fixation of favorable alleles, there was a cluster consisting of bulk breeding. No clusters formed in the extreme of the effective population size

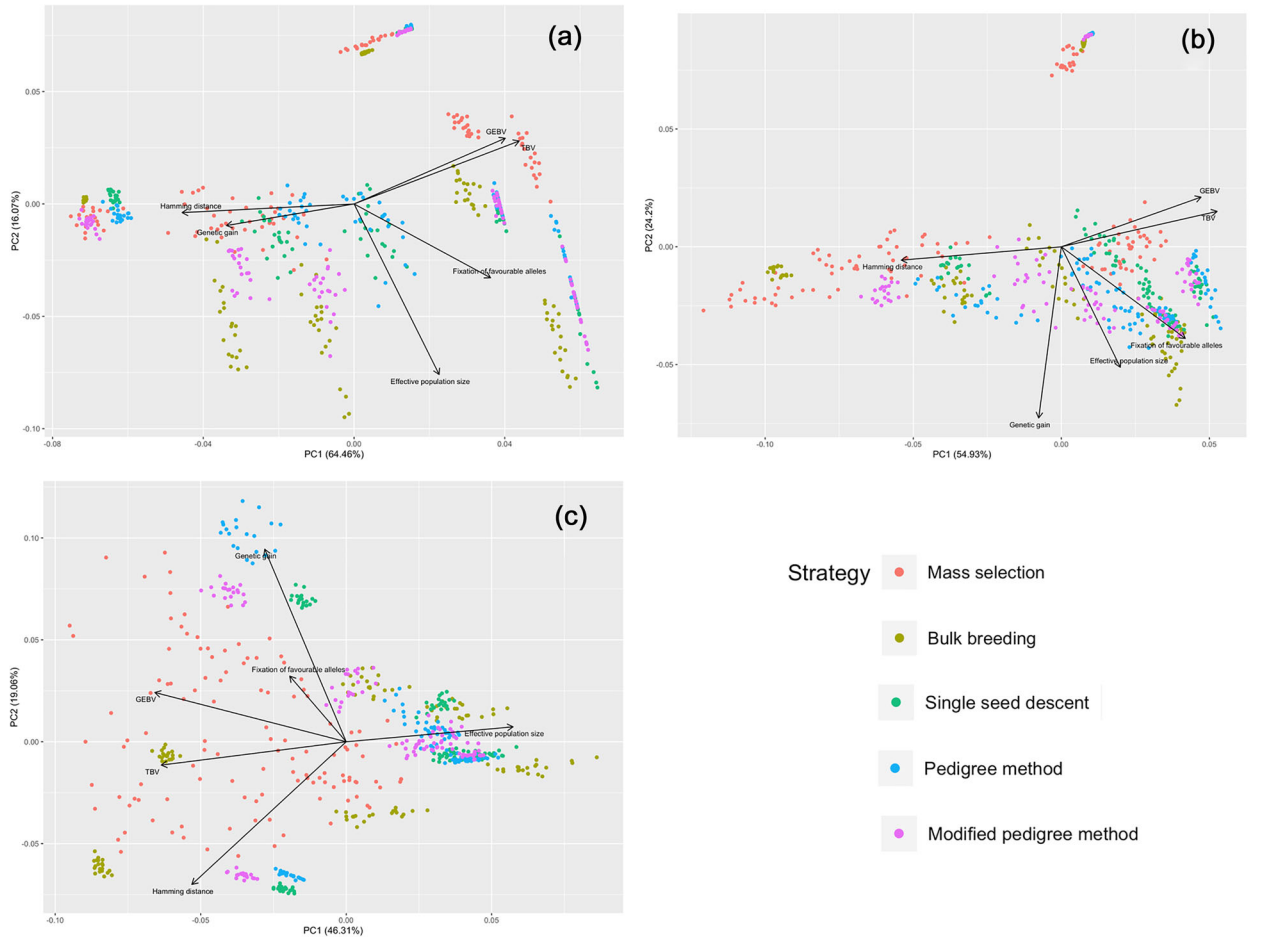

FIGURE 6 Principal component analysis (PCA) plot for (a) days to flowering, (b) white mold tolerance, and (c) seed yield following genomic selection (GS) model update. Two major principal components account for (a) 80.53%, (b) 79.13%, and (c) 65.37% of the variances, respectively. Colors correspond to the selection strategy simulated. Breeding strategies include mass selection, bulk breeding, single seed descent, the pedigree method, and the modified pedigree method.

eigenvector. Near the center of the plot was a cluster consisting of the pedigree method and single seed descent.

The first two principal components in the WM PCA (Figure 6) explained 79.13% of the variance. For DF, the eigenvectors for GEBV and TBV were close to each other. Meanwhile, the eigenvectors for Hamming distance and genetic gain were located close together. In the extreme of the Hamming distance eigenvector, there was a cluster consisting of mass selection, while for the genetic gain eigenvector, there was a cluster for the modified pedigree method. In the direction of the eigenvector for the fixation of favorable alleles, there was a cluster for bulk breeding.

For SY (Figure 6), the first two principal components described 79.47% of the variance. The GEBV and TBV eigenvectors are very close together and point in similar directions. On the opposite end are the Hamming distance and genetic gain eigenvectors, which are located close together. Toward the extreme of the Hamming distance eigenvector is a cluster made of mass selection and bulk breeding. Between the eigenvectors for fixed favorable alleles and effective population size, there was a cluster corresponding to bulk breeding.

3.6 | Genetic variance

Additive genetic variance for this simulation, evaluated in Lin et al. (2022), revealed that the relative change in variance differs between strategies and numbers of parents. To summarize, for each trait variance decreased over each cycle, and the decrease was most pronounced when the breeding program started with fewer parents. Variance under GS was equal to or greater than the variance observed in conventional breeding, especially for SY under the mass selection strategy. For DF, the most variance was conserved under bulk breeding. For WM tolerance, mass selection maintained the most variance, except when the initial parent population was the smallest (15 parents) in which case bulk breeding maintained the most variance. When conducting GS on a population with 30

parents, the modified pedigree method resulted in the greatest maintained variance, while when the population had 60 parents, mass selection resulted in the greatest maintenance of variance.

4 | DISCUSSION

The results demonstrated that prediction accuracy varied across strategy, trait, and parental population size. Genetic gain also varied along with strategy, trait, and population size. The PCAs for each trait had eigenvectors pointing in the same direction; however, the strategies did not cluster separately. This suggests that trait architecture and narrow sense heritability may have a greater influence on GS accuracy compared to other parameters. In other words, the ideal strategy and parental population size will vary by trait.

4.1 | Expected GS accuracy compared to in silico realized GS accuracy

The expected GS accuracy results suggested that increased parental population sizes should lead to a higher accuracy. In addition, this increase should be particularly evident in mass selection. However, these equation-based accuracies did not reflect the in silico realized GS accuracies estimated from TBV and GEBV correlations. Moreover, the realized accuracies fluctuated from one cycle to the next, while the predicted accuracies showed a gradual decline over the six cycles. Expected accuracies were much higher and consistent than in silico realized accuracies.

The lack of concordance between predicted and realize accuracies is likely due to our relatively small population sizes and the impact that high selection pressure has on the structure of a population. Estimates of effective population size depend on the assumption that selection is not acting on the population at hand. When artificial selection is acting on a population, the number of offspring that each parent contributes to the next generation is unequally distributed. According to Waples (2022), as the variance in the number of offspring per parent increases, N_e decreases. So, the N_e value should theoretically decrease as the population continues to undergo selection because each parent's contribution to the next generation becomes increasingly variable. Considering Equations (3) and (4), this explains the overestimations in the expected accuracies.

Brard and Ricard (2015) previously explored the ability of formulae to predict GS accuracy. They found that the method in which N_e, and by extension M_e, were estimated had an impact on how well the formulae could predict GS accuracy. Both N_e and M_e can be calculated using a number of different methods, meaning different results may be obtained even when the same formula is used to predict GS accuracy. Several models have been proposed for the estimation of N_e (Caballero & Toro, 2000; Crow & Morton, 1955; Wang & Hill, 2000; Wright, 1938). Depending on the model used, certain assumptions are made regarding the population under investigation. For plant species in particular, few estimates have been made for N_e. Siol et al. (2007) were the first to report estimates for the highly selfing model legume species, *Medicago truncatula*. Since *M. truncatula* and *Proteus vulgaris* belong to the same family *Fabaceae*, this method for estimating N_e was selected. Although the formula is not a perfect model for estimating N_e in a breeding population, it still provides us with a theoretical basis for approximating GS accuracy given the genetic makeup of our population.

Further, Equation (2) does not properly account for situations with a very large number of loci. Based on Equation (2), as the number of loci increases, the accuracy will become biased and shift toward 0. This is because there cannot be an infinite number of independent loci. Instead, independent chromosome segments were used in Equation (3) to predict GS accuracies (Daetwyler et al., 2010). Using estimates of LD as opposed to a known number of loci added a degree of abstraction to the GS accuracy estimates, possibly adding to the disparity between realized and predicted accuracies.

In the PCA plots, the eigenvectors for TBV and GEBV pointed in similar directions, which confirms that correlations are an accurate measure of GS accuracy. However, Figures 2 and 4 show the limitations to the use of correlations. In some of the later cycles, the variance was zero due to TBVs and GEBVs converging on a single value, leading to undefined correlations.

4.2 | Impact of training population size on GS accuracy

One of the major factors that influences GS accuracy is the training population size. The closed system that was simulated in QU-GENE involved taking the progeny at the end of each cycle and using them as the parents in the next cycle. Part of this process consisted of maintaining 30 families after each cycle. The SimuPOP population, which was used as the parents at the start of the first cycle, was also used for training. This resulted in a small training population. As a result, prediction accuracies were relatively low because the accuracy of the GS model increases along with the training population size (Rincent et al., 2017). There is not necessarily an ideal training population size; the number of individuals required to form a strong training population depends on the effective population size, heritability of the trait in question, the size of the genome, and the quality and volume of genotype and phenotype data (Voss-Fels et al., 2019). Although a more sizable training population may have led to greater accuracies,

we must also consider feasibility. While some bean breeding programs are large and resource rich, many are not (Beaver & Osorno, 2009), so we provide realistic estimations of GS performance even if a breeding program is on the smaller end of the spectrum. As a general rule, breeders should aim to maximize their training population size and construct a training population using genotypes that resemble the testing population.

4.3 | Impact of the selection strategy on GS accuracy

According to the PCA plots, the breeding strategies did not cluster separately from one another. Most of the clusters that formed consisted of more than one strategy. However, according to the realized accuracies, the pedigree method and single seed descent led to the greatest accuracies by the end of the 10 cycles under selection for DF and WF. Single seed descent involves advancing one seed from each plant, eliminating the impact of natural or artificial selection on genotypes. The pedigree method involves advancing seed from several superior families. Under these methods, breeders can identify and retain desirable additive effects. Further, it is possible that higher accuracies observed under these two methods are likely due to a retention of genetic variance from generation to generation. When the training population captures higher variance, the model tends to perform better on the testing population (Rincent et al., 2017). For SY, pedigree and mass selection saw the highest accuracies. The pedigree method yielded the highest accuracies with a smaller number of parents (15) supporting the idea that when there is a small base population, the pedigree method can help retain the genetic variance that was present in the training population.

4.4 | Impact of trait architecture on GS accuracies

Consistent with findings from previous studies, predictive ability for traits controlled by a small number of QTL varies depending on which prediction method is used (Wang et al., 2015). For WM tolerance and SY, realized accuracy could not be obtained for some of the later cycles. This is likely a result of certain alleles being fixed, and ultimately converging on a single breeding value. It is also plausible that more QTL are required to make accurate predictions on these traits. For SY, 11 known QTL were used, and for WM eight known QTL were used. For DF, accuracies were obtained in later cycles, and the largest number of known QTL (19) was used for this trait. When all TBVs in a cycle are the same, the correlation becomes undefined as the variance is zero. If phenotype-based and GS leads to different sets of alleles becoming fixed, it is possible that they are converging on different breeding values. This explains the poor correlation, despite TBVs and GEBVs pointing in similar directions in the PCA plots. The fixation of alleles may have also been influenced by the number of QTLs. As previously shown by Lin et al. (2022), selection for WM tolerance led to the highest percentage of fixed favorable alleles in the fewest cycles of selection, while DF had the lowest allele fixation rate. In this case, fewer simulated QTL for a given trait resulted in higher fixation, and more simulated QTL for a given trait led to lower fixation rate. This may just be due to the fact that fixation rate per total alleles will statistically be lower when more QTL are present. However, it is worth noting that DF is known to be a polygenic trait, and interactions between the many QTL controlling that trait may have also impacted the allele fixation rate (González et al., 2021). This explains why correlations could not be obtained for some of the later cycles for WM tolerance and SY. It is also important to note that for effective use of GS models, a higher number of QTLs will result in higher accuracy and higher genetic gain. GS is an effective method when selecting on quantitative traits controlled by numerous genes (Tong et al., 2021).

4.5 | Effectiveness of model updating

It is important to note that to update the model, the simulation must be stopped to generate a new model and update. Stopping and rerunning the simulation will by default result in an increase in genetic gain. To account for this, phenotype-based (or conventional) selection was included alongside GS to evaluate the impact of updating the model. Figure 5 shows that for some strategies, there was a larger increase in genetic gain when GS was implemented compared to phenotype-based selection immediately after the model update. This suggests that updating the GS model may be beneficial for certain strategies under GS. However, as was the case with the unchanged GS model, predicted accuracies did not reflect the realized accuracies when model updating was included. According to predicted accuracies, updating the GS model led to a spike in accuracy that quickly declined again in the next cycle. This spike in accuracy suggests the importance of relatedness between training and testing populations. The realized accuracies fluctuated from one cycle to the next. However, the inclusion of model-updating did lead to an increase in genetic gain.

Updating was most beneficial for the pedigree method and mass selection in the DF simulation. Model update was beneficial for the use of the pedigree method in selecting for SY. Model update did not benefit any of the strategies in WM tolerance simulation, further suggesting the significance of high quality and quantity of QTL data. Unexpectedly, the genetic gain eigenvector was closest to the Hamming distance

eigenvector and far away from the eigenvector for the fixation of favorable alleles. Thus, results for the GS model update should be taken cautiously.

For both traits, model updating had the most significant effect during the pedigree method simulations. The pedigree method leaves entire plant families behind in the selection process, so the training population for a pedigree method GS pipeline must continue to reflect the genotypes being selected on. Model updating had less of an impact on WM tolerance compared to DF and SY. Of the 1010 QTL we used to simulate common bean genotypes, only eight were known to confer WM tolerance. White mold tolerance is a complex quantitative trait (Oladzad et al., 2019), so to accurately model its genetic control more coverage of the genome would be required. The results of the model updates tell us that both relatedness and trait architecture play a crucial role in realized prediction accuracy.

4.6 | Application to common bean breeding programs

Common bean breeding programs vary widely across the globe, serving the needs of diverse customers and addressing the requirements of diverse environments. Each breeding strategy tested in this study (mass selection, bulk breeding, single seed descent, pedigree method, and modified pedigree method) will impact genetic variance and allele frequencies differently. Therefore, there will be inherent differences in GS accuracies between strategies, but our aim was to establish a generalized guide for initiating a GS in a common bean breeding program using recombinant inbred lines, as opposed to modeling one specific population structure or breeding pipeline.

If a breeder is concerned about the negative impact GS may have on the genetic variance (i.e., bottlenecks) in their population, our results indicate that mass selection or bulk breeding is the best method for maintaining genetic variance under GS. On the other hand, if genetic gain is the primary goal, our findings indicate that the pedigree method and single seed descent are most compatible with GS compared to the other strategies. One of the challenges of implementing GS in a breeding program is developing a training population that will inform our prediction model which genotypes yield desirable breeding values, and this study provided some insight into how a common bean training population should be structured. First, the training population must periodically shift as the breeding population shifts throughout breeding cycles. Second, when genotyping for GS breeders should be sure to acquire adequately dense SNP data to ensure that the model can pick up on genotype/trait associations. Poor performance due to inadequate SNP data is exemplified by the observed prediction performance for WM. SNP chips can be ideal for genotyping for GS due to the repeatability of SNPs with each retraining of the model. However, there is a limit to density with an SNP chip, so if a breeder is working on a low-heritability trait, high SNP density should be prioritized (Zhang et al., 2017). It is also worth noting that a diverse training population can make for a more comprehensive model. A diverse training population would be highly structured, with the structure depending on the gene pool (Mesoamerican or Andean), market classes, and the pedigrees of the training individuals (Amongi et al., 2023). While we provide a more generalized analysis, a breeding program wishing to implement GS should also consider the structure of their training population. Our predictions did not take into account structure, and this likely limited our prediction accuracies.

5 | CONCLUSION

GS has been widely used in animal breeding; however, its effectiveness in plant breeding still requires more validation to implement, particularly when dealing with the complex processes that take place in plant breeding program pipelines. This study showed that variation is present in genetic gain and GS accuracy when dealing with different selection strategies. Numerous studies have investigated prediction accuracy in simulations. However, in those studies, QTLs were simulated and were evenly distributed across the genome with effect sizes drawn from a random distribution. Using QTLs, effect sizes, and positions from mapping experiments, this study aimed at assessing GS accuracy in simulation that better approximates the ground truth. These simulations also reflect QTL segregating in breeding programs while modeling breeding schemes commonly used by bean breeders (Beaver & Osorno, 2009). The findings from the study indicate that predicted estimates of accuracy do not reflect of accuracies obtained from correlations between TBV and GEBV. Furthermore, according to in silico realized accuracies, there may be some benefits to using GS under single seed descent or the pedigree method. For a given parent population size, accuracies varied depending on which strategy was used, and which trait was being estimated. These findings suggest that trait architecture and breeding strategy play a more significant role in genomic prediction than the initial parent population size. When building a training population, care should be taken to select genotypes that mirror those expected in the testing population in terms of allelic effects, allele frequency, and the number of genes controlling the trait at hand. For complex quantitative traits, adequate marker density should be prioritized, and the prediction model should be updated at least once per cycle. There is no one-size-fits all solution for teaching a GS model to make accurate predictions, but simulation is an important tool for exploring the parameters of a GS model prior to implementation.

AUTHOR CONTRIBUTIONS

Isabella Chiaravallotti: Data curation; project administration; validation; writing – review & editing. **Jennifer Lin**: Data curation; formal analysis; investigation; visualization; writing – original draft. **Vivi Arief**: Methodology; software; supervision; writing – review & editing. **Zulfi Jahufer**: Software; supervision; writing – review & editing. **Juan M. Osorno**: Resources; supervision; writing – review & editing. **Phillip McClean**: Resources; writing – review & editing. **Diego Jarquin**: Writing – review & editing. **Valerio Hoyos-Villegas**: Conceptualization; data curation; funding acquisition; methodology; project administration; resources; supervision; writing – review & editing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data and codes are available at Github: https://github.com/McGillHaricots/peas-andlove/tree/master/Simulation-files.

ORCID

Juan M. Osorno https://orcid.org/0000-0003-0905-3523
Diego Jarquin https://orcid.org/0000-0002-5098-2060
Valerio Hoyos-Villegas https://orcid.org/0000-0003-1080-9148

REFERENCES

Amongi, W., Nkalubo, S. T., Ochwo-Ssemakula, M., Badji, A., Dramadri, I. O., Odongo, T. L., Nuwamanya, E., Tukamuhabwe, P., Izquierdo, P., Cichy, K., Kelly, J., & Mukankusi, C. (2023). Genetic clustering, and diversity of African panel of released common bean genotypes and breeding lines. *Genetic Resources and Crop Evolution*, *70*, 2063–2076.

Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., & Jannink, J. (2011). Accuracy and training population design for genomic selection on quantitative traits in elite North American oats. *The Plant Genome*, *4*(2). https://doi.org/10.3835/plantgenome2011.02.0007

Assefa, T., Assibi Mahama, A., Brown, A. V., Cannon, E. K. S., Rubyogo, J. C., Rao, I. M., Blair, M. W., & Cannon, S. B. (2019). A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*, *39*, 1–23. https://doi.org/10.1007/s11032-018-0920-0

Bandillo, N. B., Jarquin, D., Posadas, L. G., Lorenz, A. J., & Graef, G. L. (2023). Genomic selection performs as effectively as phenotypic selection for increasing seed yield in soybean. *The Plant Genome*, *16*(1), e20285. https://doi.org/10.1002/tpg2.20285

Beaver, J. S., & Osorno, J. M. (2009). Achievements and limitations of contemporary common bean breeding using conventional and molecular approaches. *Euphytica*, *168*, 145–175. https://doi.org/10.1007/s10681-009-9911-x

Bernardo, R. (1994). Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Science*, *34*(1), 20–25.

Bernardo, R. (2020). Reinventing quantitative genetics for plant breeding: Something old, something new, something borrowed, something BLUE. *Heredity*, *125*(6), 375–385. https://doi.org/10.1038/s41437-020-0312-1

Berro, I., Lado, B., Nalin, R. S., Quincke, M., & Gutiérrez, L. (2019). Training population optimization for genomic selection. *The Plant Genome*, *12*(3), 190028. https://doi.org/10.3835/plantgenome2019.04.0028

Beyene, Y., Semagn, K., Mugo, S., Tarekegne, A., Babu, R., Meisel, B., Sehabiague, P., Makumbi, D., Magorokosho, C., Oikeh, S., Gakunga, J., Vargas, M., Olsen, M., Prasanna, B. M., Banziger, M., & Crossa, J. (2015). Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. *Crop Science*, *55*(1), 154–163. https://doi.org/10.2135/cropsci2014.07.0460

Boopathi, N. M., & Boopathi, N. M. (2020). Marker-assisted selection (MAS). *Genetic mapping and marker assisted selection: Basics, practice and benefits* (pp. 343–388). Springer. https://doi.org/10.1007/978-981-15-2949-8_9

Brard, S., & Ricard, A. (2015). Is the use of formulae a reliable way to predict the accuracy of genomic selection? *Journal of Animal Breeding and Genetics*, *132*, 207–217. https://doi.org/10.1111/jbg.12123

Caballero, A., & Toro, M. A. (2000). Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genetical Research*, *75*, 331–343. https://doi.org/10.1017/S0016672399004449

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., De Los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic selection in plant breeding: Methods, models, and perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Crow, J. F., & Morton, N. E. (1955). Measurement of gene frequency drift in small populations. *Evolution; International Journal of Organic Evolution*, *9*, 202–214. https://doi.org/10.2307/2405589

Daetwyler, H. D., Pong-Wong, R., Villanueva, B., & Woolliams, J. A. (2010). The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, *185*, 1021–1031. https://doi.org/10.1534/genetics.110.116855

Das, R. R., Vinayan, M. T., Patel, M. B., Phagna, R. K., Singh, S. B., Shahi, J. P., Sarma, A., Barua, N. S., Babu, R., Seetharam, K., Burgueño, J. A., & Zaidi, P. H. (2020). Genetic gains with rapid-cycle genomic selection for combined drought and waterlogging tolerance in tropical maize (*Zea mays* L.). *The Plant Genome*, *13*(3), e20035.

Doublet, A.-C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., Laloë, D., & Restoux, G. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genetics Selection Evolution*, *51*, 52. https://doi.org/10.1186/s12711-019-0495-1

Dreisigacker, S., Pérez-Rodríguez, P., Crespo-Herrera, L., Bentley, A. R., & Cross, J. (2023). Results from rapid-cycle recurrent genomic selection in spring bread wheat. *G3: Genes, Genomes, Genetics*, *13*(4), jkad025.

Edwards, S. M., Buntjer, J. B., Jackson, R., Bentley, A. R., Lage, J., Byrne, E., Burt, C., Jack, P., Berry, S., Flatman, E., Poupard, B., Smith, S., Hayes, C., Gaynor, R. C., Gorjanc, G., Howell, P., Ober, E., Mackay, I. J., & Hickey, J. M. (2019). The effects of training population design on genomic prediction accuracy in wheat. *Theoretical and Applied Genetics*, *132*, 1943–1952.

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genetics*, *4*, 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Galeano, C. H., Fernandez, A. C., Franco-Herrera, N., Cichy, K., McClean, P., Vanderleyden, J., & Blair, M. (2011). Saturation of an intra-gene pool linkage map: towards a unified consensus linkage map for fine mapping and synteny analysis in common bean. *PLoS One*, *6*(12), e28135.

Gianola, D. (2013). Priors in whole-genome regression: The Bayesian alphabet returns. *Genetics*, *194*, 573–596. https://doi.org/10.1534/genetics.113.151753

González, A. M., Yuste-Lisbona, F. J., Weller, J., Vander Schoor, J. K., Lozano, R., & Santalla, M. (2021). Characterization of QTL and environmental interactions controlling flowering time in Andean common bean (*Phaseolus vulgaris* L.). *Frontiers in Plant Science*, *11*, 599462. https://doi.org/10.3389/fpls.2020.599462

Haile, J. K., N'diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., Bassi, F. M., & Pozniak, C. J. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Molecular Breeding*, *38*, 1–18. https://doi.org/10.1007/s11032-018-0818-x

Heffner, E. L., Lorenz, A. J., Jannink, J.-L., & Sorrells, M. E. (2010). Plant breeding with genomic selection: Gain per unit time and cost. *Crop Science*, *50*, 1681–1690. https://doi.org/10.2135/cropsci2009.11.0662

Hoyos-Villegas, V., Arief, V. N., Yang, W.-H., Sun, M., Delacy, I. H., Barrett, B. A., Jahufer, Z., & Basford, K. E. (2019). QuLine-Plus: Extending plant breeding strategy and genetic model simulation to cross-pollinated populations—Case studies in forage breeding. *Heredity*, *122*, 684–695. https://doi.org/10.1038/s41437-018-0156-0

Isidro, J., Jannink, J.-L., Akdemir, D., Poland, J., Heslot, N., & Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *Theoretical and Applied Genetics*, *128*, 145–158. https://doi.org/10.1007/s00122-014-2418-4

Iwata, H., Hayashi, T., & Tsumura, Y. (2011). Prospects for genomic selection in conifer breeding: A simulation study of *Cryptomeria japonica*. *Tree Genet Genomes*, *7*, 747–758. https://doi.org/10.1007/s11295-011-0371-9

Jannink, J.-L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. *Briefings in Functional Genomics*, *9*(2), 166–177. https://doi.org/10.1093/bfgp/elq001

Keller, B., Ariza-Suarez, D., De La Hoz, J., Aparicio, J. S., Portilla-Benavides, A. E., Buendia, H. F., Mayor, V. M., Studer, B., & Raatz, B. (2020). Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Frontiers in Plant Science*, *11*, 1001. https://doi.org/10.3389/fpls.2020.01001

Lin, J., Arief, V., Jahufer, Z., Osorno, J., McClean, P., Jarquin, D., & Hoyos-Villegas, V. (2022). Simulations of rate of genetic gain in dry bean breeding programs. *Theoretical and Applied Genetics*, *136*(1), 14. https://doi.org/10.21203/rs.3.rs-1442864/v1

Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., Smith, K. P., Sorrells, M. E., & Jannink, J.-L. (2011). Genomic selection in plant breeding: Knowledge and prospects. In D. L. Sparks (Ed.), *Advances in agronomy* (pp. 77–123). Academic Press.

Massman, J. M., Jung, H.-J. G., & Bernardo, R. (2013). Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. *Crop Science*, *53*, 58–66. https://doi.org/10.2135/cropsci2012.02.0112

McGill Pulse Breeding and Genetics Laboratory. (2022). *Peas-andLove*. Github. https://github.com/McGillHaricots/peas-andlove

Meuwissen, T., Hayes, B., & Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers*, *6*(1), 6–14. https://doi.org/10.2527/af.2016-0002

Meuwissen, T. H. E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution*, *41*(1), 1–9. https://doi.org/10.1186/1297-9686-41-35

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*, 1819. https://doi.org/10.1093/genetics/157.4.1819

Muleta, K. T., Pressoir, G., & Morris, G. P. (2019). Optimizing genomic selection for a sorghum breeding program in Haiti: A simulation study. *G3 Genes|Genomes|Genetics*, *9*, 391–401. https://doi.org/10.1534/g3.118.200932

Nakaya, A., & Isobe, S. N. (2012). Will genomic selection be a practical method for plant breeding? *Annals of Botany*, *110*(6), 1303–1316. https://doi.org/10.1093/aob/mcs109

Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimising genomic selection in wheat: effect of marker density, population size and population structure on prediction accuracy. *G3 Genes|Genomes|Genetics*, *8*, 2889–2899. https://doi.org/10.1534/g3.118.200311

Oladzad, A., Roy, J., Mamidi, S., Miklas, P. N., Lee, R., & McClean, P. (2019). Linked candidate genes of different functions for white mold resistance in common bean (*Phaseolus vulgaris* L) are identified by QTL-based pooled sequencing. *Frontiers in Plant Science*, *14*, 1233285.

Peng, B., & Kimmel, M. (2005). simuPOP: A forward-time population genetics simulation environment. *Bioinformatics*, *21*, 3686–3687. https://doi.org/10.1093/bioinformatics/bti584

Rincent, R., Charcosset, A., & Moreau, L. (2017). Predicting genomic selection efficiency to optimize calibration set and to assess prediction accuracy in highly structured populations. *Theoretical and Applied Genetics*, *130*, 2231–2247.

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015a). Efficient use of historical data for genomic selection: A case study of stem rust resistance in wheat. *The Plant Genome*, *8*(1), plantgenome2014–09. https://doi.org/10.3835/plantgenome2014.09.0046

Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., & Sorrells, M. E. (2015b). Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. *The Plant Genome*, *8*(2), plantgenome2014–10. https://doi.org/10.3835/plantgenome2014.10.0074

Shao, J., Hao, Y., Wang, L., Xie, Y., Zhang, H., Bai, J., Wu, J., & Fu, J. (2022). Development of a model for genomic prediction of multiple traits in common bean germplasm, based on population structure. *Plants*, *11*(10), 1298. https://doi.org/10.3390/plants11101298

Siol, M., Bonnin, I., Olivieri, I., Prosperi, J. M., & Ronfort, J. (2007). Effective population size associated with self-fertilization: Lessons from temporal changes in allele frequencies in the selfing annual Medicago truncatula. *Journal of Evolutionary Biology*, *20*, 2349–2360. https://doi.org/10.1111/j.1420-9101.2007.01409.x

Tibbs Cortes, L., Zhang, Z., & Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *The Plant Genome*, *14*, e20077. https://doi.org/10.1002/tpg2.20077

Tong, Z., Xiu, Z., Ming, Y., Fang, D., Chen, X., Hu, Y., Zhou, J., He, W., Jiao, F., Zhang, C., Shancen, Z., Jin, H., Jian, J., & Xiao, B. (2021). Quantitative trait locus mapping and genomic selection of tobacco

(*Nicotiana tabacum* L.) based on high-density genetic map. *Plant Biotechnology Reports*, *15*, 845–854.

Verges, V. L., & Van Sanford, D. A. (2020). Genomic selection at preliminary yield trial stage: Training population design to predict untested lines. *Agronomy*, *10*(1), 60. https://doi.org/10.3390/agronomy10010060

Voss-Fels, K. P., Cooper, M., & Hayes, B. J. (2019). Accelerating crop genetic gains with genomic selection. *Theoretical and Applied Genetics*, *132*, 669–686. https://doi.org/10.1007/s00122-018-3270-8

Wang, J., & Hill, W. G. (2000). Marker-assisted selection to increase effective population size by reducing mendelian segregation variance. *Genetics*, *154*, 475–489. https://doi.org/10.1093/genetics/154.1.475

Wang, X., Yang, Z., & Xu, C. (2015). A comparison of genomic selection methods for breeding value prediction. *Science Bulletin*, *60*(10), 925–935.

Waples, R. S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, *121*, 379–391. https://doi.org/10.1093/genetics/121.2.379

Waples, R. S. (2022). What is N e, anyway? *Journal of Heredity*, *113*(4), 371–379.

Wright, S. (1938). Size of population and breeding structure in relation to evolution. *Science*, *87*, 430–431. https://doi.org/10.1126/science.87.2263.425-a

Zhang, A. O., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., San Vicente, F., Olsen, M., Prasanna, B. M., Crossa, J., Yu, H., & Zhang, X. (2017). Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. *Frontiers in Plant Science*, *8*, 1916.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chiaravallotti, I., Lin, J., Arief, V., Jahufer, Z., Osorno, J. M., McClean, P., Jarquin, D., & Hoyos-Villegas, V. (2024). Simulations of multiple breeding strategy scenarios in common bean for assessing genomic selection accuracy and model updating. *The Plant Genome*, *17,* e20388. https://doi.org/10.1002/tpg2.20388