The Impact of Genetic Architecture on Genome-wide Evaluation Methods

Hans D. Daetwyler*,§, Ricardo Pong-Wong*, Beatriz Villanueva**,§§, John A. Woolliams*

*The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin, UK, EH25 9PS; §Animal Breeding and Genomics Centre, Wageningen University, 6700 AH Wageningen, NL; **Scottish Agriculture College, Edinburgh, UK, EH9 3JG; §§Departamento de Mejora Genética Animal, INIA (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria), 28040 Madrid, Spain

Running head: Impact of Genetic Architecture

Key words: Genomic Structure, Genetic Architecture, Genome-wide Evaluation, Quantitative Trait, Independent Chromosome Segments

Corresponding author: Hans D. Daetwyler, Bioscience Research Division, Department of Primary Industries, 1 Park Drive, Bundoora, 3083, Victoria, Australia. Phone: +61 3 9479 5387, Fax: +61 3 9479 3618, hans.daetwyler@dpi.vic.gov.au
The rapid increase in high throughput single nucleotide polymorphism data have led to a great interest in applying genome-wide evaluation methods to identify an individual’s genetic merit. Genome-wide evaluation combines statistical methods with genomic data to predict genetic values for complex traits. Considerable uncertainty currently exists in determining which genome-wide evaluation method is the most appropriate. We hypothesise that genome-wide methods deal differently with the genetic architecture of quantitative traits and genomes. A genomic linear method (GBLUP), and a genomic non-linear Bayesian variable selection method (BayesB) are compared using stochastic simulation across three effective population sizes and a wide range of numbers of quantitative trait loci ($N_{QTL}$). GBLUP had a constant accuracy, for a given heritability and sample size, regardless of $N_{QTL}$. BayesB had a higher accuracy than GBLUP when $N_{QTL}$ was low, but this advantage diminished as $N_{QTL}$ increased and when $N_{QTL}$ became large, GBLUP slightly outperformed BayesB. In addition, deterministic equations are extended to predict the accuracy of both methods and to estimate the number of independent chromosome segments ($M_e$) and $N_{QTL}$. The predictions of accuracy and estimates of $M_e$ and $N_{QTL}$ were generally in good agreement with results from simulated data. We conclude that the relative accuracy of GBLUP and BayesB for a given number of records and heritability, are highly dependent on $M_e$, which is a property of the target genome, as well as the architecture of the trait ($N_{QTL}$).
The rapid progress and reducing costs of genome sequencing and high throughput DNA-techniques have led to a great interest in applying genome-wide evaluation methods to identify individuals of high genetic merit. Genome-wide evaluation uses associations of a large number of SNP (single nucleotide polymorphisms) markers across the whole genome with phenotypes to produce accurate estimates of breeding values (EBVs) for candidates to selection (Meuwissen et al. 2001). The accuracy of genome-wide selection (i.e., selection based on genomic EBVs) is expected to be substantially higher than that of traditional BLUP (best linear unbiased prediction) selection which is based on pedigree and phenotypic data (Daetwyler et al. 2008; Goddard 2008; Hayes et al. 2009c). In addition, GWS has the potential to reduce inbreeding rates because of the increased emphasis on own rather than family information (Woolliiams et al. 2002; Dekkers 2007; Daetwyler et al. 2007). Furthermore, the application of genome-wide evaluation approaches can significantly aid our understanding of quantitative trait genetic architecture.

The genome-wide evaluation methods suggested to date can be broadly categorized into groups according to whether or not there is assortment of the SNP by magnitude of effect or contribution to the variance. One group treats SNP homogeneously and includes variants of genomic best linear unbiased prediction (GBLUP). This group includes a form of ridge regression (Meuwissen et al. 2001) and the use of a realised relationship matrix computed from the markers instead of the traditional pedigree matrix (Villanueva et al. 2005; NejatiJavaremi et al. 1997; Hayes et al. 2009c). Both approaches have been shown to be equivalent (Habier et al. 2007; Goddard 2009). A second group provides for heterogeneity among SNP contributions to the variance, with some contributions permitted to be large whilst the remainder are small, possibly zero. This assortment is helped by Bayesian approaches which place priors on numbers of SNP with major contributions (e.g. BayesA,
BayesB, Meuwissen et al. 2001; Lee et al. 2008; Meuwissen et al. 2009), or with some penalty based on functions of the magnitude of effect for each SNP (e.g. Lasso, Tibshirani 1996; Yi & Xu 2008) or with other smoothing metrics (Long et al. 2007). A third group attempts to reduce dimensionality by using principal components or partial least squares (Raadsma et al. 2008; Solberg et al. 2009) to identify an informative subset of SNP genotypes. The main two methods currently used in real datasets are a linear prediction method, GBLUP, and variants of non-linear Bayesian variable selection approaches such as BayesB.

In most simulated published data, the accuracy of BayesB outperformed that of GBLUP (e.g. Meuwissen et al. 2001; Habier et al. 2007; Lund et al. 2009). However, real data results have not consistently supported this conclusion. Two reviews of empirical results in dairy cattle to date have shown that GBLUP and BayesB result in very similar accuracies for most traits (VanRaden et al. 2009; Hayes et al. 2009a). One reason for the disagreement between simulated and real data results could be that the genetic architecture simulated is significantly different from what is found in real populations. Most studies published to date that compare methods using simulated architectures have considered only 50 or fewer QTL affecting the trait (e.g. Meuwissen et al. 2001; Habier et al. 2007; Lund et al. 2009). In this paper we hypothesise that the relative utility of genome-wide evaluation methods depends significantly on both the genomic structure of the population and the genetic trait architecture.

The main objective of this study was to compare a linear method, GBLUP, and a non-linear variable selection method, BayesB, using simulated data across a range of population and trait genetic architectures to further understand the mechanics of genome-wide evaluation methods. An important secondary objective was to extend deterministic prediction models to predict the accuracy of both methods. Theoretical models complement stochastic simulation
by helping the understanding of the factors involved in genome-wide evaluation performance
and, in return, stochastic simulation is used to confirm theoretical derivations.

METHODS

**Theoretical development:** Daetwyler et al. (2008) derived equations for predicting the
accuracy of a simple least-squares genome-wide evaluation approach for continuous and
dichotomous traits. The original formula for genome-wide accuracy for a continuous trait is
\[ r_{\hat{a}} = \sqrt{\left( \frac{N_P h^2}{N_P h^2 + n_G} \right)} \]
where \( r_{\hat{a}} \) is the correlation between true and estimated
additive genetic values (i.e. accuracy), \( N_P \) is the number of individuals in the training
population, \( h^2 \) is the heritability and \( n_G \) is the number of independent loci (Daetwyler et al.
2008). The accuracy was independent of how large the subset of loci was that make non-zero
contributions. Thus, it did not matter whether there were many non-zero loci effects of small
magnitude or only a few non-zero loci effects of large magnitude. In Daetwyler et al. (2008),
the formulae were derived by considering the regression of phenotypes on one locus at a time
as loci were assumed independent. Therefore the formula will work for small numbers of
dispersed loci in a genome but the accuracy will tend to zero as \( n_G \) becomes large;
erroneously, because loci cannot be added independently in a finite genome due to linkage.
Daetwyler et al. (2008) discussed that an empirical value for the number of independent
chromosome segments (\( \hat{M}_L \)) could be used in place of \( n_G \), because \( n_G \) was assumed
independent. Goddard (2009) also proposed accuracy predictions for GBLUP which used the
concept of predicted \( M_c \). His derivation builds on work by Visscher et al. (2006) in which
the variance of identical-by-descent sharing for full sibs was developed, and provides a
prediction for \( M_c \) which is \( M_c = \frac{2N_c L}{\log(4N_c L)} \), where \( L \) is the genome length in
Morgans (Goddard 2009). Substituting $M_e$ in place of $n_C$ in the original formula of Daetwyler et al. (2008) results in:

$$r_{\theta \theta G} = \sqrt{\frac{N_P h^2}{N_P h^2 + M_e}}, \quad (1)$$

which also predicts the accuracy of GBLUP. At no time does the argument moving from the original formula of Daetwyler et al. (2008) to Equation (1) depend on the distribution of the effects of loci, so we come to the first hypothesis in this study which states that GBLUP accuracy is independent of the number of quantitative trait loci ($N_{QTL}$) associated with the phenotypic trait.

Our second hypothesis was that the accuracy of BayesB when $N_{QTL}$ is high would tend to that of GBLUP. If our first hypothesis is confirmed, then the dependence of GBLUP on $M_e$ is an advantage at high $N_{QTL}$, even though $N_{QTL}$ may be higher than $M_e$. Heuristically, if GBLUP delivers accuracy as if there are a $M_e$ number of QTL, the benefit from prior information that there are approximately $M_e$ (or more) QTL is unclear given equation (1). On the other hand, it is a clear disadvantage if $N_{QTL} < M_e$ because GBLUP cannot adapt the model to suit the data. In contrast, BayesB is a variable selection method which attempts to determine the ‘optimum dimensionality’ given the data and prior information. When $N_{QTL}$ is high this optimum is likely to be $M_e$ in both methods. Hence, the accuracy of BayesB at high $N_{QTL}$ can be predicted in the same way as GBLUP, but if $N_{QTL} < M_e$, variable selection may deliver an advantage in accuracy because choosing a subset of variables will reduce the dimensionality of the model. Thus, substituting $N_{QTL}$ for $M_e$ is likely to better predict the accuracy of BayesB. This results in the following equation,

$$r_{\theta \theta B} = \sqrt{\frac{N_P h^2}{N_P h^2 + \min(N_{QTL}, M_e)}}. \quad (2)$$
Further rearrangement of Equations (1) and (2) allows for empirical estimates of $\hat{M}_e$ ($\hat{M}_e$) to be made in the following way,

$$\hat{M}_e = \frac{(N_P h^2)(1 - r_{\theta\theta}^2)}{r_{\theta\theta}^2}$$  \hspace{1cm} (3)

where $r_{\theta\theta}^2$ is the squared accuracy of estimates of genetic values using GBLUP or BayesB (when $\hat{N}_{QTL} \geq \hat{M}_e$) for individuals without phenotypes. Predicting $\hat{M}_e$ with GBLUP requires molecular relatedness to be known, whereas this is not required when using BayesB. This result gives a further sub-hypothesis that the empirical $\hat{M}_e$ is predicted by the formula for independent chromosome segments given by Goddard (2009). Also, if $N_{QTL} < M_e$, additional information on $\hat{N}_{QTL}$ can be gathered using BayesB accuracy because it can choose a subset of loci or variables, by applying the following formula:

$$\hat{N}_{QTL} = \frac{(N_P h^2)(1 - r_{\theta\theta}^B)}{r_{\theta\theta}^B}$$  \hspace{1cm} (4)

Therefore, additional insight into quantitative traits can be gained by combining genome-wide evaluation and deterministic prediction. The accuracy of BayesB is of course influenced by the priors used in the analyses (especially priors on the proportion of loci with no effect) hence it is important to use appropriate priors to get accurate $\hat{N}_{QTL}$.

**Simulations**

Our study consisted of three main steps. First, populations of individuals were simulated to be in mutation drift equilibrium. Second, effects were assigned to a number of QTL that were randomly selected from the whole set of segregating loci and true genetic values and phenotypes were generated for each individual. The third step consisted of the genetic evaluations of the individuals generated with both GBLUP and BayesB.

**Populations and Genome:** Populations in mutation drift equilibrium were simulated by random mating individuals for many generations with recombination and mutation. The number of male and female parents was $\frac{1}{2} N_e$ across generations. A total of 1,000, 5,000 and
10,000 generations were simulated until linkage disequilibrium and heterozygosity values were stable for \( N_e = 200 \), \( N_r = 1,000 \) and \( N_r = 2,000 \), respectively. In the final generation, a set of training individuals (of variable size) in which the loci effects were to be estimated was generated by random mating. Using the same parents, a set of validation individuals of size equal to the training set was produced whose genetic values were to be predicted. In scenarios where the size of the training sets \( (N_p) \) was larger than \( N_e \), population size was increased by increasing the number of offspring per mating in the final generation. The total genome size was 10 Morgans (10 chromosomes of 1 Morgan each). In generation zero all individuals were completely homozygous for the same allele and mutations were applied at a rate of \( 2.5 \times 10^{-5} \) per locus per meiosis in the following generations. Mutations switched allele one to two and vice versa. The number of randomly distributed mutations per chromosome was sampled from a Poisson distribution with mean corresponding to the product of the number of loci per chromosome and the mutation rate. Similarly, recombinations per chromosome were sampled from a Poisson distribution with a mean of one per M and were then randomly placed along the chromosome. Linkage disequilibrium (LD) statistics; i.e., \( R^2 \) (Hill & Robertson 1968) between adjacent segregating loci were averaged among all pairs exceeding a minor allele frequency of 0.05 and matched expected \( R^2 \) values (Sved 1971; Tenesa et al. 2007). Allele frequency distributions were found to follow a U-shaped distribution.

The number of loci at the start of the simulation (generation zero) required several considerations concerned with obtaining appropriate number of segregating loci \( (N_L) \) and \( N_{OTL} \) in the final generation. The realised relationship matrix used in GBLUP can be singular if \( N_L \) is less than the number of individuals in the matrix (VanRaden 2008), preventing the inversion needed to compute solutions. Thus, \( N_L \) at mutation drift equilibrium was made larger than the maximum sum of training and validation individuals to be used, and a similar
was used across all scenarios. However, as \( N_e \) increased the proportion of segregating loci in the last generation also increased and for \( N_e = 200, N_e = 1,000, \) and \( N_e = 2,000, \) approximately 0.04, 0.28 and 0.52 of initial loci were segregating at mutation drift equilibrium, respectively. This required the adjustment of the number of initial loci. Across all scenarios, \( N_L \) varied between 4576 and 4721 loci (SE < 9.5 loci). To obtain \( N_{QTL} M_e \) in a random mating population, as derived by Goddard (2009), was used as a guide to allow comparisons to be made across \( N_e. \) The following \( N_{QTL} \) scenarios were simulated: 0.03, 0.05, 0.30, 0.50, 0.75 and 1\( M_e. \) Table 1 outlines the corresponding \( N_{QTL} \) for these proportions of \( M_e \) for the three \( N_e, \) whereas Table 2 shows all scenarios which were carried out. Note that throughout this study our use of the terms ‘low’ or ‘high’ \( N_{QTL} \) may refer to different actual \( N_{QTL} \) across the three \( N_e \) because \( N_{QTL} \) was scaled to be proportional to \( M_e \) (Table 1).

The desired \( N_{QTL} \) were randomly chosen from \( N_L. \) True allele substitution effects were sampled from \( N(0,1). \) True breeding values for 2\( N_p \) (i.e. training and validation set) individuals were calculated for each QTL as \( 2(1 - p_j)\beta_j \) (where \( p_j \) is the major allele frequency at locus \( j), -2p_j\beta_j \) and \( ((1 - p_j) - p_j)\beta_j \) for the major and minor homozygote and heterozygote genotype, respectively (Falconer & Mackay 1996). Total breeding values were obtained by summing over \( N_{QTL}. \) Breeding values were scaled to have the variance of \( h^2 \) (i.e. phenotypic variance = 1). Phenotypic records were simulated for \( N_p \) (training set) animals by adding independent environmental terms drawn from \( N(0,1-h^2) \) to true breeding values. In addition, sampling \( \beta_j \) from a Laplace (double exponential) distribution was investigated in Scenario 4. The accuracy of both genome-wide evaluation methods was computed as the correlation between true and estimated breeding values.

**GBLUP analysis:** The evaluation with GBLUP applied the following model which was fit in ASReml (Gilmour et al. 1995): \( y = \mu 1 + Za + e, \) where \( y \) is the vector of phenotypic
values, $\mu$ is the population mean, $Z$ is an incidence matrix for random individual effects, $a$ is a vector of random individual additive genetic values and $e$ is the residual. Random effects $a$ and $e$ were assumed normally distributed as $N(0, G\sigma_a^2)$ and $N(0, I\sigma_e^2)$, respectively, where $G$ was the realised relationship matrix computed using the $N_I$ loci. In $G$, the relationship between a pair of individuals was based on identical-by-state probabilities and included all training individuals with phenotypes and validation individuals without phenotypes. The total allelic relationship at a locus between a pair of individuals was calculated as $0.5 \sum_{i=1}^{2} \sum_{j=1}^{2} \delta_{ij}$, where $\delta_{ij}$ is 1 if allele $i$ in the first individual is identical to allele $j$ in the second individual and 0 otherwise. Averaging over loci as $0.5 \sum_{i=1}^{2} \sum_{j=1}^{2} \delta_{ij} [N_L]^{-1}$ yields the numerator relationship between all individual pairs required for $G$ (NejatiJavaremi et al. 1997). Breeding values were obtained by solving the mixed model equations (Henderson 1975).

**BayesB analysis:** We implemented a variant of the original BayesB (Meuwissen et al. 2001).

The model applied was $y = \mu + \sum_{j=1}^{N_L} X_j \beta_j + e$, where $\mu$ was the mean, $X$ was an incidence equal to -1, 0, and 1 for 11, 12/21, and 22, respectively, $\beta$ is the allele substitution effect for locus $j$ and $e$ was the vector of residuals distributed as $N(0, \sigma_e^2)$. Allele substitution effects were assumed to come for a mixture distribution where a proportion ($\pi$) had no effect and a proportion $(1 - \pi)$ had effects distributed as $N(0, \sigma_{snp}^2)$. The prior used for $\mu$ was uniform, over a long range (i.e. $-100\sigma_p^2$ to $+100\sigma_p^2$), and for $\sigma_e^2$ was uniform (i.e. 0 to $+100\sigma_p^2$), where $\sigma_p^2$ is the phenotypic variance of the dataset analysed. The prior for $\sigma_{snp}^2$ was taken from a scaled inverted chi square distribution of the form $\sigma_{snp}^2 \sim \nu s^2 \chi^2_{\nu}$, where $s^2$ represented the prior value for $\sigma_{snp}^2$ and $\nu$ was its degree of belief (Wang et al. 1994). A weak prior was chosen to be 1 for both $\nu$ and $s^2$ across all scenarios. The influence of our $s^2$ on $\sigma_{snp}^2$ estimates from BayesB was investigated and was found to be
small. A critical discussion of the impact of priors on BayesB can be found in (Gianola et al. 2009).

In order to test effect of $\pi$ on breeding value accuracy, two types of scenarios were carried out. In the majority of Scenarios, $1 - \pi$ was fixed to true values simulated for a particular Scenario, $1 - \pi = N_{\text{QTL}}[N_L]^{-1}$, which we call the ‘informed prior’. However, in Scenario 8 $1 - \pi$ was fixed to 57 QTL for all $N_{\text{QTL}}$ Scenarios, $1 - \pi = 57 \text{QTL}[N_L]^{-1}$, which we term the low prior. While not done in this study, one could estimate $1 - \pi$ instead of using a prior and this may be an advantage in real data with unknown $N_{\text{QTL}}$ (e.g. Pong-Wong & Hadjipavlou 2009).

Our implementation of sampling loci variances was slightly different from Meuwissen et al. (2001). They performed a haplotype analysis and therefore several haplotype effects would need estimating per locus and having more than one effect per locus required sampling of locus variances. In this study, biallelic loci were simulated and, because only one effect was estimated per locus, it was not necessary to sample a per locus variance. This process of sampling from one variance is more similar to Xu (2003). However, his implementation assumed all loci had an effect, which is comparable to BayesA of Meuwissen et al. (2001). In contrast, our model assumed that some loci will not have an effect ($\pi$), which is more similar to BayesB. In addition, while Xu (2003) used a naïve prior for $\sigma^2_{\text{SNP}}$, we used an informative prior (i.e. $\nu$ and $\kappa^2 = 1$) as proposed by ter Braak et al. (2005).

The length of the Gibbs chain was 105,000 iterations and the first 5,000 iterations were discarded as burn-in. Estimates at every 20th iteration were stored as a sample resulting in a total of 5,000 samples. Several analyses were carried out to test if the sampling protocol was appropriate. Autocorrelations of effects and estimated breeding values were found to be close to zero in stored samples which showed that they were almost independent, so Monte Carlo error was < 0.0003 in all scenarios (Geyer & Thompson 1992). This allowed for
shortening of the chain length to 45,000 iterations (2000 samples) for scenarios with \( N_P = 2,000 \) to reduce running time. Furthermore, convergence was investigated by using a variety of starting values for \( \sigma^2 \) in the BayesB. Using different starting values resulted in nearly identical estimates of \( \sigma^2_{\text{snps}} \) and breeding value accuracy suggesting that our Gibbs chains were converging. In addition, a long chain of 160,000 iterations was run and \( \sigma^2_{\text{snps}} \) estimates and accuracy values were very similar to the shorter chain. While convergence can not be guaranteed in any MCMC study, evidence suggests that our MCMC protocol converged.

RESULTS

**GBLUP accuracy:** The accuracy of GBLUP for a given set of values for \( N_P \) and \( h^2 \) stayed constant regardless of \( N_{\text{OTL}} \) in all scenarios simulated (Figure 1 and Table 3). This confirmed our first hypothesis. The constant accuracy results from the unique \( M_e \) of a random mating population which, in turn, depends on \( N_e \) and \( L \) (Goddard 2009). The plateau of GBLUP accuracy increased when more phenotypic records were used in the estimation of breeding values and when \( h^2 \) increased (Figure 1). Sampling \( \beta_i \) from a Laplace (double exponential) distribution resulted in the same GBLUP accuracy as sampling from \( N(0, 1) \).

**BayesB accuracy:** In contrast to GBLUP, with BayesB the accuracy was highest at low \( N_{\text{OTL}} \) and then decreased as \( N_{\text{OTL}} \) increased (Figure 1 and Table 3). Once \( N_{\text{OTL}} \) was high BayesB reached a plateau where the accuracy does not decrease anymore despite increasing \( N_{\text{QTL}} \). This plateau was observed in all BayesB scenarios and the value of the accuracy at this plateau depended on \( N_e, h^2 \) and \( N_P \) (Tables 3 and 4). The plateau decreased when \( N_e \) increased. An increase in \( h^2 \) and \( N_P \) influenced the accuracy in two ways, firstly, it raised the overall accuracy in all \( N_{\text{OTL}} \) scenarios. Secondly, it slightly shifted the onset of the accuracy plateau to higher \( N_{\text{QTL}} \). Sampling effects from a Laplace distribution instead of \( N(0, 1) \)
raised BayesB accuracy slightly, but when \( N_{\text{QTL}} \) was equal to 0.03 \( M_e \) or \( N_{\text{QTL}} > 0.5 M_e \) no difference in accuracy was observed between sampling from both distributions.

The use of low \( N_{\text{QTL}} \) priors for \( 1 - \pi \) yielded a lower accuracy than informed priors (Scenario 8, Table 2; Figure 2). The gap between the accuracy of informed and low priors increased as \( N_{\text{QTL}} \) increased because the proportion of the genetic variance explained by the low \( N_{\text{QTL}} \) prior became smaller.

**Comparison of GBLUP and BayesB:** The comparison of GBLUP and BayesB leads to several key observations. BayesB always performed better than GBLUP at low \( N_{\text{QTL}} \). However, as \( N_{\text{QTL}} \) increased, the difference between the two methods became smaller and eventually both approaches achieved very similar accuracy. The \( N_{\text{QTL}} \) at which this equivalence occurred was increased with increasing \( N_e \), \( N_p \) and \( h^2 \) (Figure 1, Tables 3 and 4). Once \( N_{\text{QTL}} \) increased past the equivalence point, BayesB had a slightly lower accuracy than GBLUP and settled at a constant accuracy (Table 3). The difference between GBLUP and BayesB at high \( N_{\text{QTL}} \) decreased when \( N_p \) was increased. Sampling effects from a Laplace instead of a Normal distribution did not affect these general trends.

In Figure 2, the maximum x-value of \( N_{\text{QTL}} \) plotted is equal to the predicted \( M_e \) from Goddard (2009) and one observes that BayesB accuracy approaches the plateau and becomes similar to GBLUP accuracy well below \( M_e \). A first inspection therefore suggests that the second hypothesis does not hold. However the argument for the second hypothesis is based upon the empirical \( \hat{M}_e \). \( \hat{M}_e \) calculated with Equation (3) for \( h^2 = 0.1, 0.3 \) and 0.5 by averaging over the values of \( N_{\text{QTL}} \) gives values of 890, 900 and 700. In this context, hypothesis two is shown to be broadly valid, in that superiority of BayesB over GBLUP disappears when \( N_{\text{QTL}} \) approaches \( \hat{M}_e \), although there is a trend for BayesB accuracy to
become similar to GBLUP accuracy sooner than $M_e$. These observations held for other scenarios. The comparison between $M_e$ and $\hat{M}_e$ is addressed in more detail below.

**Decay in accuracy:** The decay in accuracy between training and validation individuals was also greater for GBLUP than that of BayesB at low $N_{QTL}$ (Table 4) as observed by Habier et al. (2007). However, this trend diminished as $N_{QTL}$ increased and the decay of accuracy reached similar levels in both methods at high $N_{QTL}$.

**Predictions of accuracy:** Figure 3 shows the accuracies of GBLUP and BayesB predicted with Equations (1) and (2), respectively and the accuracies from simulations in the validation set. Predictions of GBLUP and BayesB (at high $N_{QTL}$) accuracy were generally accurate. The accuracy of the predictions were highly dependent on $M_e$. In BayesB, the drop in accuracy as $N_{QTL}$ increased was predicted well. Equation (2) tended to over-predict BayesB accuracy, particularly in scenarios where $N_{QTL}$ was a low proportion of $M_e$, using Goddard (2009), and low $h^2$ and $N_p$.
Empirical $\hat{M}_e$ and $\hat{N}_{QTL}$: We estimated $M_e$ using the accuracy of GBLUP or BayesB (when

$N_{QTL} = M_e$) (Equation (3)). When GBLUP accuracy was used, we averaged the accuracy

across all $N_{QTL}$ scenarios simulated for a given set of values for $h^2$ and $N_P$. This was done

for each population replicate to obtain a standard error. It is a sub-hypothesis that $M_e$ as

predicted by Goddard (2009) approximates $\hat{M}_e$. Empirical estimates of $M_e$ using GBLUP

were always lower than those using BayesB (Table 5) due to the higher GBLUP accuracy

when $N_{QTL}$ is high. The estimates using BayesB accuracy were more variable than GBLUP

as shown by the larger SE of BayesB. A general trend was apparent showing that $\hat{M}_e$

increased as $N_P$ increased which suggests that $\hat{M}_e$ has not reached a bound, however the

change in $\hat{M}_e$ is small in relation to the difference from $M_e$. Furthermore, $\hat{M}_e$ does not

increase linearly with $N_P$ and this may indicate that it may be approaching asymptotic values.

The number of QTL controlling the trait ($N_{QTL}$) was estimated using Equation (4) with

reliability values from BayesB when $N_{QTL} < M_e$. As shown in Figure 4 for Scenario 7, the

estimated $N_{QTL}$ do follow the actual $N_{QTL}$ well and are predictive of the trend. Empirical

$\hat{N}_{QTL}$ were better estimated with higher $N_P$. Note that incorrect priors will reduce

accuracy.

DISCUSSION

We have compared GBLUP and BayesB at various population and trait genetic architectures

and at various $N_P$. We demonstrated that GBLUP had a constant accuracy, for a given $N_P$

and $h^2$, regardless of $N_{QTL}$. The accuracy of BayesB was greatest at low $N_{QTL}$, decreased

with increasing $N_{QTL}$ and eventually reached a lower accuracy plateau below which the

accuracy did not fall even when $N_{QTL}$ was further increased. BayesB has an advantage over

GBLUP at low $N_{QTL}$, but this advantage decreased as $N_{QTL}$ increased and it finally
diminished completely or, in some cases, the advantage switched to GBLUP depending on $N_c$ and $N_P$. The point at which GBLUP and BayesB accuracy became equal was related to the empirical number of independent segments estimated from the GBLUP accuracy, $\hat{M}_c$, which was less than the theoretical prediction of $M_c$ provided by Goddard (2009). It is clear from this study that quantifying the superiority of GBLUP over BayesB or vice versa depends upon three sets of attributes: the population genome structure (e.g. $N_c$), the trait genetic architecture (e.g. $N_{QTL}$, $h^2$) and the size of the training set. Superiority is, therefore, not a property of the method and general statements to that effect should be avoided. Furthermore, we have proposed and tested equations for the prediction of GBLUP and BayesB accuracy and the estimation of $\hat{M}_c$ and $\hat{N}_{QTL}$. Our predictions follow achieved GBLUP and BayesB accuracy well. Empirical $\hat{M}_c$ values seem to be approaching an asymptote with increasing $N_P$, and our estimates of $\hat{N}_{QTL}$ follow the trend of true $N_{QTL}$. The constant accuracy of GBLUP, for a given $h^2$ and $N_P$, confirmed our first hypothesis and clearly shows that this accuracy depends crucially on genomic properties, and not on properties of the trait, and is summarised in the concept of $M_c$. In turn, $M_c$ will depend on $N_c$ and $T$, which can be viewed, in the short term, as constants in a random mating population (Goddard 2009). In a wider sense, $M_c$ and the more commonly known haplotype blocks are both measures resulting from population history and as such are related, although not interchangeable. Haplotype blocks are physical segments of the genome within which haplotype diversity is low, bounded by areas where evidence for historical recombination exists (e.g. Goldstein 2001; Gabriel et al. 2002; Frazer et al. 2007). While haplotype blocks are a physical measure, $M_c$ is a more statistical concept related to the behaviour of the genome in genomic evaluations. It is theoretically derived from variation in relationship between relatives and from the continuum of variation in linkage disequilibrium across the genome (Goddard 2009). Both the number of haplotype blocks and $M_c$ increase with
increasing $N_e$ and in close relatives haplotype blocks will be long and $M_e$ will be small (Hayes et al. 2009c). It should be noted that the dependence of GBLUP on $M_e$, shown in this study does not support the conclusion that GBLUP assumes an infinitesimal model in which there are a very large number of genes each contributing a small portion to the genetic variance. In fact, GBLUP is indifferent to $N_{QTL}$ unless $N_{QTL}$ is very small (i.e. one or $<10$ QTL, unpublished results), as demonstrated in this study.

While it is clear that in GBLUP the accuracy depends on $M_e$ and not on $N_{QTL}$, in BayesB the accuracy depends on the interplay of both features of genetic architecture ($N_{QTL}$ and $M_e$). Our results follow our second hypothesis that the behaviour of BayesB accuracy at high $N_{QTL}$ is similar to that of GBLUP. The accuracy of BayesB declines as $N_{QTL}$ increases and eventually becomes similar to GBLUP as $N_{QTL}$ increases. The point at which this occurs approaches $M_e$ with increasing $N_P$. Therefore, we show that the accuracy of BayesB at high $N_{QTL}$ is also dependent on $M_e$ just like in GBLUP. This is also supported by the accuracy plateau being observed across at similar proportions of $M_e$ and that near equivalence is approximated closely by $N_{QTL} = \hat{M}_e$, where $\hat{M}_e$ is the empirical estimate of $M_e$ obtained from GBLUP. Therefore, the plateau is not a function of actual $N_{QTL}$ but of $M_e$. Another argument for $\hat{M}_e$ to be a major determinant in BayesB accuracy at high $N_{QTL}$ is that it can be accurately predicted with Equation (3). In addition, a recent study of Barley using real data has also shown that GBLUP accuracy exceeded that of BayesB when $N_{QTL}$ was large (Zhong et al. 2009).

The difference in accuracy between GBLUP and BayesB may be explained by the way both methods process and model information. With GBLUP, each independent segment is estimated irrespective of whether it has an effect or not whereas with BayesB, an additional step is involved in which for each locus it is estimated if the locus has an effect or not. The fact of choosing a near correct subset of loci with effect with BayesB increases accuracy but
there is also an error associated with determining $\pi$ which depends on $N_L$. When $N_{QTL} < M$ the advantage of choosing a subset is clear but when $N_{QTL} \geq M$ this diminishes (heuristically, it is likely that each independent segment contains QTL). Thus, under the latter scenario, GBLUP performs slightly better than BayesB. This argument is further supported by the decreasing difference in accuracy between GBLUP and the BayesB accuracy plateau at high $N_{QTL}$ when $N_P$ is increased, because BayesB can identify the loci with effect better with more information.

The use of QTL effects sampled from a Laplace distribution instead of a Normal distribution resulted in no quantitative change to GBLUP accuracy and in only minor changes in accuracy to BayesB. Overall, the trends observed were fully consistent with the conclusions obtained with normally distributed effects, confirming other studies which have compared these effect distributions (Habier et al. 2007; Daetwyler et al. 2008; Meuwissen 2009).

It is challenging to compare statistical approaches that embody different statistical philosophies. In the genomic and animal breeding context, comparing approaches using point estimates of breeding values, such as their accuracy, is common. Breeding itself is an application of decision theory (e.g. Woolliams et al. 1993) and in model terms the value of a breeding decision may be measured as the expected rate of genetic gain ($\Delta G$) arising from the decision. The focus on accuracy in this paper arises from the ‘breeders equation’ $E[\Delta G] = ir\sigma_A/L$, where $i$ is standardised intensity, $\sigma_A$ is the additive genetic standard deviation, $L$ is generation interval and $r$ is the accuracy. The concept of accuracy is therefore central to the infrastructure of genetic evaluation.

This does not mean that accuracy as presented here will remain the key comparison in future, and the richer information available from Bayesian methods may at some point overturn this paradigm. The expectation of the posterior distribution for an individual’s breeding value is a natural estimate arising from a Bayesian analysis (e.g. Goddard, 2009; Meuwissen et al. 2009).
and also gives a natural comparison with the estimates from BLUP approaches. In addition, alternative measures of comparison to expected gain (in a Bayesian sense, minimising the squared loss) have been considered, such as percentiles of the posterior which will be influenced by (co)variances of candidates (e.g. Woolliams et al. 1993), or other approaches to selection such as minimax regret.

Genome-wide evaluation methods are popular because they seem to offer a solution to predicting a large numbers of parameters (i.e. loci effects) from a limited number of phenotypes, but a number of concerns have been expressed. For example, Bayesian methods without an influential prior may experience convergence problems. Investigation with non-informative priors for scale factors have resulted in similar BayesB accuracy to those using priors in this study (unpublished results). However, more investigation of convergence of genome-wide methods is needed, including minimum length of Gibbs chains, influence of priors on convergence and optimal protocols for parallel computing. At the same time, there is concern that the priors on degree of belief and scale factors used in the BayesB (Meuwissen et al. 2001) exert too much influence thereby preventing Bayesian learning (Gianola et al. 2009). Gianola et al. (2009) clearly show that priors influence estimates.

The findings that both GBLUP and BayesB depend significantly on $M_s$ are given more weight by the fact that the accuracy of both methods can be predicted with Equations (1) and (2), respectively. The predictions were generally accurate but limitations have also been highlighted, especially in predicting BayesB accuracy. Extensions to the formulae may be needed to predict BayesB more accurately at low $N_{OTL}$ relative to $M_s$ when $h^2$ or $N_D$ are also low, and there is also a need to review whether $M_s$ as formulated by Goddard (2009) is a good predictor of $M_e$. However, being able to predict the trend in BayesB accuracy is a significant step forward (Figure 3). One of the assumptions in the original derivation (Daetwyler et al. 2008) was that all of the genetic variance was tagged by the loci used in the
analysis. This represents a complication when applying our equations to predict the accuracy using a commercially available SNP chip, because the current chips are likely to miss a portion of the genetic variance. Firstly, it is likely that the number of SNP on current chips is not high enough to tag all the genetic variance and variation not associated with SNP (e.g. copy number variation, Redon et al. 2006) will also be missed. Secondly, SNP with higher than average heterozygosity are selected for developing the chips and therefore loci with low minor allele frequency are proportionally underrepresented (i.e. ascertainment bias). The result of this missing genetic variance in the analysis of real populations is that our deterministic equations are likely to over-predict the accuracy in both methods.

The fact that our equations account for the entire genetic variance will, however, be a clear advantage as the scientific community moves towards the analysis of sequence data for which our formulae are appropriate in their current form. In sequence data analysis, all base-pairs are included and therefore no rare alleles would be missing. Thus, all the genetic variance is contained in the sequence and the prediction does not rely on capturing LD with the true mutation.

Additional insight into quantitative traits can be gained by combining genome-wide evaluation and deterministic prediction. We have shown that $M_r$ can be estimated with Equation (3) if the accuracy of GBLUP or BayesB is known. Two theoretical values for $M_e$ have been proposed to date, $M_e = 2N_eL/\log(\Lambda N_eL)$ (Goddard 2009) and $2N_eL$ (Hayes et al. 2009c). Our estimates of $M_r$ ($\hat{M}_r$), even though still increasing with increasing $N_p$, remain lower than both theoretical values but were of the right order of magnitude when using Goddard’s equation (2009) rather than $2N_eL$ (Table 5). In real data using $2N_eL$ in Equation (1) appears to predict GBLUP accuracy well (Hayes et al. 2009b), but this may be due to the fact that SNP arrays miss a significant proportion of the genetic variance. Once more of the genetic variance is captured with new technology we would expect that estimates of $\hat{M}_r$ from
real data would likely tend towards the derivation of Goddard (2009), or possibly lower. In addition to \( \hat{M}_c \), \( N_{\text{QTL}} \) can be estimated with BayesB accuracy if \( N_{\text{OTL}} < M_c \). As Figure 4 shows, this can be a coarse measure of \( N_{\text{QTL}} \), because small changes in accuracy can cause relatively large fluctuations in \( \hat{N}_{\text{OTL}} \). A complication in estimating \( N_{\text{OTL}} \) is that several SNP may be in partial LD with a particular QTL and this could lead to overestimates of \( N_{\text{OTL}} \). In addition, BayesB requires knowledge of the true \( N_{\text{QTL}} \) in its prior. Nevertheless, estimates of \( N_{\text{QTL}} \) could aid investigations into complex trait architectures, perhaps through examining the correspondence between the assumed prior on \( N_{\text{OTL}} \) and the resulting estimate.

The trends observed in this study are supported by experiences in real data. Results in dairy cattle genotyped with a 50K SNP chip show that GBLUP and BayesB lead to very similar accuracies in most traits (VanRaden et al. 2009; Hayes et al. 2009a). VanRaden et al. (2009) report correlations between linear and nonlinear methods of >0.99 in a vast majority of traits. This suggests that in real animal populations quantitative traits are controlled by a large number of QTL and for most traits \( N_{\text{OTL}} \geq \hat{M}_c \). There are of course exceptions to the rule and, for example, in dairy cattle BayesB performed better than GBLUP in milk fat content (VanRaden et al. 2009). This is likely due to a significant portion of the variation being explained by few genes of large effect, such as DGAT (Grisart et al. 2004). Hence, in this trait it is likely that \( N_{\text{QTL}} < M_c \) or that a relatively small number of QTL explain the majority of the genetic variance in the trait. We have investigated a scenario where one QTL explained 25% of the genetic variance and 1000 very small QTL explained the rest of the variance (results not shown) and the results confirmed our hypothesis.

The principles established in this study should be transferable to other populations as the trends have been confirmed across three different \( N_c \). In our view, investigators need to gather evidence to answer two questions. Firstly, what is the population’s \( M_c \) and, secondly, how many \( \hat{N}_{\text{OTL}} \) are likely contributing to the genetic variance in a particular trait. When
$N_{\text{QTL}} \geq M_e$ GBLUP will result in higher accuracy than BayesB, but when $N_{\text{QTL}} < M_e$ BayesB will outperform GBLUP.

ACKNOWLEDGEMENTS

We are grateful to Piter Bijma, Jesús Fernández and two anonymous reviewers for their helpful and constructive comments. HDD was supported by the SABRETRAIN Project, which is funded by the Marie Curie Host Fellowships for Early Stage Research Training, as part of the 6th Framework Programme of the European Commission. BV received support from the Scottish Executive Environment and Rural Affairs Department (SEERAD) and INIA, and JAW received funding from the Biotechnology and Biological Sciences Research Council (BBSRC). This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF, http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk).


Meuwissen, T. H. E., 2009 Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution* 41.


TABLE 1. - Number of QTL simulated for each proportion of independent chromosome segments ($M_e$) for three values of effective population size ($N_e$).

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>0.03$M_e$</th>
<th>0.05$M_e$</th>
<th>0.15$M_e$</th>
<th>0.3$M_e$</th>
<th>0.5$M_e$</th>
<th>0.75$M_e$</th>
<th>1$M_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>13</td>
<td>22</td>
<td>67</td>
<td>133</td>
<td>223</td>
<td>334</td>
<td>445</td>
</tr>
<tr>
<td>1000</td>
<td>57</td>
<td>94</td>
<td>283</td>
<td>566</td>
<td>944</td>
<td>1415</td>
<td>1887</td>
</tr>
<tr>
<td>2000</td>
<td>106</td>
<td>177</td>
<td>531</td>
<td>1063</td>
<td>1772</td>
<td>2657</td>
<td>3543</td>
</tr>
</tbody>
</table>
TABLE 2. - Parameter values for the simulated scenarios, where $N_e$ = effective population size; $N_p$ = number of individuals in the training set; $h^2$ = heritability; Prior = prior used for $\pi$ in BayesB; $M_c$ = number of independent chromosome segments.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$N_e$</th>
<th>$N_p$</th>
<th>$h^2$</th>
<th>Prior</th>
<th>$M_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>200</td>
<td>0.3</td>
<td>Informed</td>
<td>445</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>1000</td>
<td>0.3</td>
<td>Informed</td>
<td>445</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>1000</td>
<td>0.5</td>
<td>Informed</td>
<td>1887</td>
</tr>
<tr>
<td>4</td>
<td>1000</td>
<td>1000</td>
<td>0.3</td>
<td>Informed</td>
<td>1887</td>
</tr>
<tr>
<td>5</td>
<td>1000</td>
<td>1000</td>
<td>0.1</td>
<td>Informed</td>
<td>1887</td>
</tr>
<tr>
<td>6</td>
<td>1000</td>
<td>500</td>
<td>0.3</td>
<td>Informed</td>
<td>1887</td>
</tr>
<tr>
<td>7</td>
<td>1000</td>
<td>2000</td>
<td>0.3</td>
<td>Informed</td>
<td>1887</td>
</tr>
<tr>
<td>8</td>
<td>1000</td>
<td>1000</td>
<td>0.3</td>
<td>57 QTL</td>
<td>1887</td>
</tr>
<tr>
<td>9</td>
<td>2000</td>
<td>2000</td>
<td>0.3</td>
<td>Informed</td>
<td>3543</td>
</tr>
</tbody>
</table>
TABLE 3. - Accuracy of GBLUP and BayesB (informed priors for \( \pi \)) for different effective population sizes \((N_e)\), numbers of QTL expressed as proportions of \( M_e \), numbers of individuals in the training set \((N_P)\) when the heritability is 0.3. SE < 0.023 in all scenarios.

<table>
<thead>
<tr>
<th>Method</th>
<th>(N_e)</th>
<th>( N_P)</th>
<th>(0.03M_e)</th>
<th>(0.05M_e)</th>
<th>(0.15M_e)</th>
<th>(0.3M_e)</th>
<th>(0.50M_e)</th>
<th>(0.75M_e)</th>
<th>(1M_e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBLUP</td>
<td>200</td>
<td>200</td>
<td>0.405</td>
<td>0.450</td>
<td>0.429</td>
<td>0.414</td>
<td>0.444</td>
<td>0.416</td>
<td>0.398</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1000</td>
<td>0.505</td>
<td>0.501</td>
<td>0.508</td>
<td>0.502</td>
<td>0.507</td>
<td>0.501</td>
<td>0.511</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>2000</td>
<td>0.575</td>
<td>0.579</td>
<td>0.571</td>
<td>0.568</td>
<td>0.571</td>
<td>0.571</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>200</td>
<td>0.739</td>
<td>0.649</td>
<td>0.463</td>
<td>0.400</td>
<td>0.398</td>
<td>0.365</td>
<td>0.344</td>
</tr>
<tr>
<td>BayesB</td>
<td>1000</td>
<td>1000</td>
<td>0.865</td>
<td>0.772</td>
<td>0.601</td>
<td>0.516</td>
<td>0.480</td>
<td>0.451</td>
<td>0.445</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>2000</td>
<td>0.886</td>
<td>0.812</td>
<td>0.646</td>
<td>0.573</td>
<td>0.544</td>
<td>0.522</td>
<td>0.506</td>
</tr>
</tbody>
</table>

29
TABLE 4. - Accuracy of GBLUP and BayesB (informed priors for $\pi$) in training (T) and validation (V) individuals for different effective population sizes ($N_e$), numbers of QTL ($N_{QTL}$) expressed as proportions of $M_e$ and numbers of individuals in the training set ($N_P$) when the heritability is 0.3. SE < 0.023 in all scenarios.

<table>
<thead>
<tr>
<th>$N_{QTL}$</th>
<th>Pop.</th>
<th>$N_P$ = 200</th>
<th>$N_P$ = 1000</th>
<th>$N_P$ = 500</th>
<th>$N_P$ = 1000</th>
<th>$N_P$ = 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>0.794</td>
<td>0.958</td>
<td>0.801</td>
<td>0.883</td>
<td>0.933</td>
</tr>
<tr>
<td>0.03$M_e$</td>
<td>V</td>
<td>0.739</td>
<td>0.952</td>
<td>0.757</td>
<td>0.865</td>
<td>0.924</td>
</tr>
<tr>
<td>0.05$M_e$</td>
<td>V</td>
<td>0.649</td>
<td>0.905</td>
<td>0.602</td>
<td>0.772</td>
<td>0.870</td>
</tr>
<tr>
<td>0.15$M_e$</td>
<td>V</td>
<td>0.463</td>
<td>0.803</td>
<td>0.421</td>
<td>0.601</td>
<td>0.744</td>
</tr>
<tr>
<td>0.30$M_e$</td>
<td>V</td>
<td>0.400</td>
<td>0.709</td>
<td>0.371</td>
<td>0.516</td>
<td>0.656</td>
</tr>
<tr>
<td>0.50$M_e$</td>
<td>V</td>
<td>0.398</td>
<td>0.654</td>
<td>0.373</td>
<td>0.480</td>
<td>0.613</td>
</tr>
<tr>
<td>1$M_e$</td>
<td>V</td>
<td>0.344</td>
<td>0.591</td>
<td>0.342</td>
<td>0.445</td>
<td>0.567</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0.599</td>
<td>0.778</td>
<td>0.583</td>
<td>0.657</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>0.344</td>
<td>0.591</td>
<td>0.342</td>
<td>0.445</td>
<td>0.567</td>
</tr>
</tbody>
</table>

GBLUP

All T 0.656 0.771 0.625 0.682 0.747

V 0.444 0.622 0.407 0.507 0.612
TABLE 5. - Estimated ($\hat{M}_e$) and predicted ($M_e$ and $M_{eh}$) number of independent chromosome segments. Estimates were obtained from Equation (3) with mean squared accuracy of GBLUP or BayesB from 50 replicates of simulated data (+SE). The number of QTL is 1 and predictions are $M_e = 2N_eL/log(4N_eL)$ as in Goddard (Goddard 2009) and $M_{eh} = 2N_eL$ as in Hayes et al. (Hayes et al. 2009c).

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>$N_p$</th>
<th>$h^2$</th>
<th>GBLUP</th>
<th>BayesB</th>
<th>$M_e$</th>
<th>$M_{eh}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>200</td>
<td>0.3</td>
<td>294 ±29</td>
<td>579 ±73</td>
<td>445</td>
<td>4,000</td>
</tr>
<tr>
<td>200</td>
<td>1000</td>
<td>0.3</td>
<td>487 ±17</td>
<td>584 ±25</td>
<td>445</td>
<td>4,000</td>
</tr>
<tr>
<td>1000</td>
<td>1000</td>
<td>0.5</td>
<td>883 ±24</td>
<td>1103 ±27</td>
<td>1887</td>
<td>20,000</td>
</tr>
<tr>
<td>1000</td>
<td>1000</td>
<td>0.3</td>
<td>890 ±28</td>
<td>1243 ±33</td>
<td>1887</td>
<td>20,000</td>
</tr>
<tr>
<td>1000</td>
<td>1000</td>
<td>0.1</td>
<td>822 ±41</td>
<td>1551 ±111</td>
<td>1887</td>
<td>20,000</td>
</tr>
<tr>
<td>1000</td>
<td>500</td>
<td>0.3</td>
<td>803 ±39</td>
<td>1187 ±48</td>
<td>1887</td>
<td>20,000</td>
</tr>
<tr>
<td>1000</td>
<td>2000</td>
<td>0.3</td>
<td>1014 ±21</td>
<td>1280 ±26</td>
<td>1887</td>
<td>20,000</td>
</tr>
<tr>
<td>2000</td>
<td>2000</td>
<td>0.3</td>
<td>1253 ±24</td>
<td>1769 ±42</td>
<td>3543</td>
<td>40,000</td>
</tr>
</tbody>
</table>
FIGURE 1. - Accuracy of GBLUP and BayesB (informed priors for $\pi$) in validation individuals for different numbers of QTL and heritabilities ($h^2$) when the effective population size is 1,000 and the number of individuals in the training set is 1,000. SE < 0.018 in all scenarios.

![Graph showing accuracy of GBLUP and BayesB for different numbers of QTL and heritabilities](image-url)
FIGURE 2. Accuracy of BayesB in validation individuals with informed priors on $\pi$ (BayesB Prior Informed) and a low prior of 57 QTL regardless of the actual number of QTL (BayesB Prior 57 QTL) when the effective population size is 1,000, the number of individuals in the training set is 1,000, and heritability is 0.3. Accuracy of GBLUP is included for reference. SE < 0.009 in all scenarios.
FIGURE 3. - Predicted (black bars) and simulated (grey bars) accuracy of GBLUP and BayesB for a) a heritability ($h^2$) of 0.3 and varying effective population size ($N_e$) and number of individuals in the training set ($N_p$) and for b) $N_e = 1000$ and varying $h^2$ and $N_p$. Different numbers of QTL expressed as proportions of $M_e$ were considered for BayesB.
FIGURE 4. - Actual number of QTL simulated and number of QTL predicted with Equation (4) using BayesB accuracy when the effective population size is 1,000 and the number of individuals in the training set is 2,000 and the heritability is 0.3.