IJAS Iranian Journal of Applied Animal Science

Examining Different Models of Gene Action in Genomic Evaluation

Research Article

H. Sahebalam[1*], M. Gholizadeh[1] and S.H. Hafezian[1]

[1] Department of Animal Science, Faculty of Animal and Aquatic Science, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

Received on: 31 Dec 2024
Revised on: 6 Feb 2025
Accepted on: 21 Feb 2025
Online Published on: Jun 2025

*Correspondence E-mail: hamid.sahebalam@yahoo.com

Online version is available on: www.ijas.ir
https://doi.org/10.71798/ijas.2025.1214798

ABSTRACT

The purpose of this study was to compare different models of gene action in genomic selection using two statistical methods including Genomic Best Linear Unbiased Prediction (GBLUP) and Bayesian Least Absolute Shrinkage and Selection Operator (BLASSO). Therefore, three gene action models including purely additive effects (A), additive effects plus dominant deviations (AD) and additive effects plus dominant deviations plus epistasis interactions (ADE) were fitted to the data by each method (i.e., GBLUP or BLASSO, GBLUP-D or BLASSO-D and GBLUP-DE or BLASSO-DE, respectively). Real genotypic data of mice were used and, phenotypes were simulated from these real genotypes based on different number of quantitative trait loci (QTLs) and levels of broad-sense heritability (*T1*:20 QTLs and H^2=0.4, *T2*:20 QTLs and H^2=0.8, *T3*: 100 QTLs and H^2=0.4, *T4*:100 QTLs and H^2=0.8, *T5*:200 QTLs and H^2=0.4, *T6*:200 QTLs and H^2=0.8). BLASSO recorded a higher prediction accuracy (PA) than GBLUP. This increase in PA was greater when the number of QTLs was low (20 QTLs), and this advantage decreased with the increase in the number of QTLs. When traits were under the control of a small number of QTLs, the PA of model A was slightly higher than AD and ADE models. The GBLUP-D or BLASSO-D methods showed a lower bias compared to GBLUP or BLASSO methods, and this bias was minimized in GBLUP-DE or BLASSO-DE methods for all traits.

KEY WORDS bias, dominance deviation, epistasis interaction, GEBV, GEGV, prediction accuracy.

INTRODUCTION

Genomic selection (GS) based on genome-wide single nucleotide polymorphism (SNP) markers refers to the prediction of breeding values and the subsequent selection of individuals (Meuwissen *et al.* 2001). The genome-wide best linear unbiased prediction (GBLUP) method utilizes genomic information in the form of a genomic relationship matrix, which defines the additive genetic covariance between individuals (VanRaden, 2008; Hayes *et al.* 2009a). Since genomic information captures Mendelian sampling across the genome, covariance between individuals can be estimated with higher accuracy compared to traditional pedigree-based methods. GBLUP has become a widely used approach in GS for dairy cattle (McHugh *et al.* 2011; Wiggans *et al.* 2011) due to its simplicity and relatively low computational requirements (Hayes *et al.* 2009b; VanRaden *et al.* 2009).

Since additive genetic merit is directly transmitted to the next generation, genetic evaluations typically focus on predicting the additive value of alleles while ignoring non-additive effects. Dominance, a source of non-additive genetic variation, refers to interactions between alleles at a single locus, whereas interactions between alleles at different loci are referred to as epistasis. When non-additive genetic effects contribute significantly to variation, relying solely on additive effects or breeding values (BV) may result in the selection of genotypes that do not possess the

highest genetic potential. Genomic information provides new opportunities for estimating non-additive genetic effects by enabling the estimation of an individual's total genetic value, leading to improved phenotype predictions. While dominance effects can be estimated alongside additive effects and summarized across loci, the estimation of epistatic effects is computationally challenging due to the exponential increase in dimensionality (Toro and Varona, 2010; Vitezica *et al.* 2013).

Over the past decade, GS has become a standard in the genetic evaluation of livestock populations. However, most GS implementations consider only additive effects when calculating the genomic estimated breeding value (GEBV) for selection candidates. Despite this, the inclusion of non-additive effects is of interest because: (i) they can improve the prediction accuracy (PA) of GEBVs, thereby enhancing selection response (Toro and Varona, 2010; Aliloo *et al.* 2016; Duenk *et al.* 2017); (ii) they facilitate mate allocation strategies for selection candidates (Mäki-Tanila, 2007; Toro and Varona, 2010; Aliloo *et al.* 2017); and (iii) they can be leveraged to optimize non-additive genetic variation through the implementation of appropriate crossbreeding or purebred breeding schemes (Mäki-Tanila, 2007; Zeng *et al.* 2013).

Non-additive effects have traditionally been excluded from genetic evaluation models due to factors such as the lack of informative pedigrees and computational challenges (Varona *et al.* 2018). While additive variance accounts for some dominant biological effects of genes (Hill, 2010), dominant deviation variance should not be overlooked.

The availability of large genomic datasets has revived interest in incorporating non-additive genetic effects into genomic models. Several researchers have proposed different approaches to account for dominance effects in genomic prediction models (Toro and Varona, 2010; Su *et al.* 2012; Vitezica *et al.* 2013).

In a simulated study with purely additive effects, the prediction accuracy (PA) of the GBLUP and BayesC methods was evaluated under different genetic architectures including different levels for the number of QTLs and the distribution of QTL effects (Shirali *et al.* 2015). Also, the PA of frequentist and Bayesian methods was compared in the analysis of simulated traits with purely additive effects (Sahebalam *et al.* 2024a). While previous studies have focused on evaluating the performance of GBLUP and BLASSO, most have either ignored or inadequately incorporated non-additive genetic effects (Hayes *et al.* 2010; Atefi *et al.* 2016; Sahebalam *et al.* 2022). Furthermore, while the effects of heritability and the number of quantitative trait loci (QTLs) on PA have been well documented (Daetwyler *et al.* 2010; Combs and Bernardo, 2013), few studies have explored how these factors interact with gene action models that account for both additive and non-additive effects. Additionally, much of the existing research has been limited to simulations with simplified genetic architectures, often overlooking real-world complexities in genetic variation (Resende *et al.* 2012; Sahebalam *et al.* 2019).

The objective of this study was to investigate and evaluate different models of gene action, including: (i) purely additive effects, (ii) additive effects plus dominance deviations, and (iii) additive effects plus dominance deviations plus epistatic interactions in genomic prediction. These models were analyzed using relationship matrix-based methods such as GBLUP, as well as marker effect-based methods such as Bayesian Least Absolute Shrinkage and Selection Operator (BLASSO).

MATERIALS AND METHODS

Mice data

The dataset contains 1,814 individuals, each genotyped for 10,346 polymorphic markers. It can be accessed at this link. The dataset originates from an experiment designed to identify quantitative trait loci (QTL) associated with complex traits in mouse populations (Valdar *et al.* 2006a; Valdar *et al.* 2006b). These data have previously been analyzed to compare genome-wide genetic evaluation methods (Legarra *et al.* 2008).

Simulated phenotypes

Phenotypes were simulated using real genotypes. A total of 20, 100, and 200 markers were randomly sampled as QTLs to control the genetic effects of the simulated phenotypes. The genetic effects were defined as the sum of all QTL effects. Three models of gene action were used for phenotype simulation: purely additive effects (A), additive effects plus dominance deviations (AD), and additive effects plus dominance deviations plus epistatic interactions (ADE).

The additive effect (a) of a QTL was defined as half the difference in genotypic values between the alternate homozygotes and was sampled from a standard normal distribution. Dominance effects (d) were defined as the deviation of the heterozygote's genotypic value from the mean genotypic values of the two homozygotes. To simulate dominance effects, dominance degrees (h_k) were first sampled from a normal distribution (N(0.5,1)). Then, absolute dominance effects were calculated as:

$$d_k = h_k \,.\, |a_k|$$

Where:

$|a_k|$: absolute value of the additive effect of the *k*th QTL.

Epistatic effects followed a gamma distribution with shape and scale parameters of 0.1 and 10, respectively (Abdollahi *et al.* 2020). Epistasis was modeled only between pairs of QTLs and included four types of interactions: additive × additive (A×A), additive × dominance (A×D), dominance × additive (D×A), and dominance × dominance (D×D). Each QTL interacted with two adjacent QTLs. Residual effects followed a normal distribution with mean 0 and variance:

$$\left((1-h^2)\sigma_g^2\right)/h^2$$

Where:
σ_g^2: represents the genotypic variance.
h^2: heritability.

After accounting for the observed genetic variance, residual effects were added to the genetic effects to generate the final phenotypes.

Model A

$$y_i = \sum_{k}^{nQTL} X_{ik} a_k + e_i$$

Where:
X_{ik} (i=1,…,n and k=1,…,nQTL): an element of the design matrix for additive genetic effects (a_k).
e_i: residual random effect.

Genotypes are coded as 2 for A_1A_1, 1 for A_1A_2, and 0 for A_2A_2 to capture additive effects.

Model AD

$$y_i = \sum_{k}^{nQTL} (X_{ik} a_k + D_{ik} d_k) + e_i$$

Where:
D_{ik} (i=1,…,n and k=1,…,nQTL): an element of the design matrix for dominance genetic effects (d_k).

Dominance effects are coded as 0, 1, and 0 for genotypes A_1A_1, A_1A_2, and A_2A_2 , respectively.

Model ADE

$$y_i = \sum_{k}^{nQTL} X_{ik} a_k + \sum_{k}^{nQTL} D_{ik} d_k + \sum_{k=1}^{nQTL-1} \sum_{k'=2}^{nQTL} l_k l'_k aa + \sum_{k=1}^{nQTL-1} \sum_{k'=2}^{nQTL} l_k l'_k ad + \sum_{k=1}^{nQTL-1} \sum_{k'}^{nQTL} l_k l'_k da + \sum_{k=1}^{nQTL-1} \sum_{k'=2}^{nQTL} l_k l'_k dd + e_i$$

Where:
$l_k l'_k aa$ (coded as 0, 1, 2 and 4), $l_k l'_k ad$ (coded as 0, 1 and 2), $l_k l'_k da$ (coded as 0, 1 and 2), and $l_k l'_k dd$ (coded as 0 and 1): epistasis effects of A×A, A×D, D×A, and D×D between two QTL k and k´, respectively.

For the A model, narrow-sense heritability (h^2) was set to 0.15 or 0.3. In contrast, for the AD and ADE models, broad-sense heritability (H^2) was set to 0.2 or 0.4, and 0.4 or 0.8, respectively.

In the A model, the variance of the residual distribution was adjusted so that the ratio of additive genetic variance to phenotypic variance equaled narrow-sense heritability. For models incorporating non-additive effects (AD and ADE), the variance of the residual distribution was set so that the ratio of total genetic variance (including additive and non-additive components) to phenotypic variance equaled broad-sense heritability.

In this study, six traits were simulated, each varying in the number of controlling loci and heritability levels (Table 1).

Table 1 Number of controlling loci and broad-sense heritability (H^2) for the six simulated traits (T1 to T6)[1]

H^2	Number of QTLs		
	20	100	200
0.4	T1	T3	T5
0.8	T2	T4	T6

[1] T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.

Statistical analysis

GBLUP

In this method, the standard BLUP mixed model equation is used, but instead of the inverse of the pedigree relationship matrix (A^{-1}), the inverse of the genomic relationship matrix (G^{-1}) is applied (Habier *et al.* 2007; Hayes *et al.* 2009b). It is assumed that marker effects follow a normal distribution.

BLASSO

In the Bayesian Least Absolute Shrinkage and Selection Operator (BLASSO) method (Park and Casella, 2008), marker effects follow a double-exponential distribution. This distribution, similar to the t-distribution, has a long tail but retains many small nonzero effects. The double-exponential distribution can also be expressed as a normal distribution, where the variance is sampled from an exponential distribution:

$$P(\beta_j \mid \lambda, \sigma_e) = DE\left(\beta_j \mid 0, \frac{\lambda}{\sigma_e^2}\right)$$

Where:

β_j: *j*th marker effect.

σ_e^2: residual variance.

The prior marginal distribution of the regression coefficients is as follows:

$$P(\beta_j|\lambda)=\int N(\beta_j|0, \sigma_e^2 \tau_j^2)\, EXP(\tau_j^2|\lambda^2)\, d\tau_j^2$$

This relation corresponds to a double-exponential distribution, which exhibits a higher density at zero, causing a sharp drop for markers with relatively small effects, while allowing a slower decline for markers with larger effects. In this model, λ is assumed to be unknown and follows a gamma distribution with shape parameter λ and scale parameter τ. BLASSO is a specialized form of penalized least squares, minimizing the sum of squared residuals. In this approach, the prior distribution of the variance of the marker effects is defined as follows:

$$Var(\beta_j|\lambda^2, \sigma_e^2) = \sigma_\beta^2 = \frac{2\sigma_e^2}{\lambda^2}$$

Regarding: $\frac{\sigma_\beta^2}{\sigma_e^2} = \frac{2}{\lambda^2}$, λ can be calculated as:

$$\sqrt{\frac{2(1-h^2)}{h^2}\left(n^{-1}\sum_i\sum_j x_{ij}^2\right)}.$$

Where:

σ^2_β: marker effects.

h^2: heritability.

x_{ij}: *j*th marker genotype in the *i*th individual.

Additive, dominance and epistatic relationship matrices

The additive genomic relationship matrix G was calculated as follows (VanRaden, 2008):

$$G = \frac{ZZ'}{2\sum_{j=1}^{m} p_j(1-p_j)}$$

Where:

Z: matrix with dimensions of n × p (the number of individuals × the number of markers) and is centered by M-P; where M is the genotype matrix coded as 0, 1 and 2 according to the number of alternative alleles; P is the matrix of locus scores $2p_j$ where p_j being the *j*th alternative allele frequency and $2\sum_{j=1}^{m} p_j(1-p_j)$ is the additive variance of markers summed across loci.

The dominance deviation genomic relationship matrix D was estimated as follows (Vitezica *et al.* 2013):

$$D = \frac{WW'}{\sum_{j=1}^{m}\left(2p_j(1-p_j)\right)^2}$$

Where:

W: same dimension as in Z, with elements equal to $-2q_j^2$, $2p_j q_j$ and $-2p_j^2$ for alternative homozygote, heterozygote and reference allele homozygote of the *j*th marker, respectively.

$\sum_{j=1}^{m}\left(2p_j(1-p_j)\right)^2$: dominance variance of markers summed across loci.

The epistasis genomic relationship matrix *I* was estimated as follows (Vitezica *et al.* 2018):

The epistasis matrices were calculated based on the two-order epistatic interaction by the Hadamard product, which is cell-by-cell multiplication and, trace defined to be the sum of the elements of the main diagonal. The epistasis matrices were calculated based on the two-locus model as $I_{AA} = \frac{G \odot G}{tr(G \odot G)/n}$, $I_{AD} = \frac{G \odot D}{tr(G \odot D)/n}$, $I_{DA} = \frac{D \odot G}{tr(D \odot G)/n}$, and $I_{DD} = \frac{D \odot D}{tr(D \odot D)/n}$, for additive × additive (A×A), additive × dominance (A×D), dominance × additive (D×A), and dominance × dominance (D×D) terms, respectively.

Different models of gene activity

Model including purely additive effects (A)

The model was used to run GBLUP, and was as follows:

$$y = Xb + Tu + e$$

For the BLASSO model:

$y = Xb + Zg_a + e$

Where:
y: vector of phenotypic observations.
b: vector of fixed effect (overall mean).
u: vector of additive genetic effects which follows a normal distribution with expectation of $N(0, G\sigma_a^2)$ where G was described earlier and σ_a^2 is the additive genetic variance.
g_a: vector of marker effects; $e \sim N(0, I\sigma_e^2)$ is the vector of random residual effects where I denotes the identity matrix and σ_e^2 is the residual variance.
X, T and Z: incidence matrices for b, u and g_a, respectively.

Model including additive effects plus dominance deviations (AD)
The model was used to run GBLUP-D, and was as follows:

$y = Xb + Tu + Td + e$

For the BLASSO-D model:

$y = Xb + Zg_a + Wg_d + e$

Where:
d: vector of dominance genetic effects which follows a normal e.g., $d \sim N(0, D\sigma_d^2)$ where D was described above and σ_d^2 is the dominance genetic variance.
g_d: vector of dominance marker effects.
W: incidence matrix for g_d.

Model including additive effects plus dominance deviations plus epistatic interactions (ADE)
The model was used to run GBLUP-DE, and was as follows:

$y = Xb + Tu + Td + Ti_{aa} + Ti_{ad} + Ti_{da} + Ti_{dd} + e$

For the BLASSO-DE model:

$y = Xb + Zg_a + Wg_d + E_1g_{aa} + E_2g_{ad} + E_2g_{da} + E_3g_{dd} + e$

Where:
i_{aa}, i_{ad}, i_{da} and i_{dd}: vectors of additive × additive (A×A), additive × dominance (A×D), dominance × additive (D×A), and dominance × dominance (D×D) epistatic effects, which are assumed to follow normal distributions with expectations $\sim N(0, I_{AA}\sigma_{aa}^2)$, $\sim N(0, I_{AD}\sigma_{ad}^2)$, $\sim N(0, I_{DA}\sigma_{da}^2)$ and $\sim N(0, I_{DD}\sigma_{dd}^2)$, respectively, where I_{AA}, I_{AD}, I_{DA} and I_{DD} were described above and σ_{aa}^2, σ_{ad}^2, σ_{da}^2 and σ_{dd}^2 are the epistatic interaction variances.
T: incidence matrix for vectors of additive, dominance and epistasis genetic effects.
g_{aa}, g_{ad}, g_{da} and g_{dd}: vectors of epistasis marker effects and E_1, E_2, E_2 and E_3 are the incidence matrices for g_{aa}, g_{ad}, g_{da} and g_{dd}, respectively.

Implementation of statistical methods
The different gene action models were implemented using the GBLUP and BLASSO methods via the BGLR package (Perez and de los Campos, 2014) in R software. A Monte Carlo Markov Chain (MCMC) sampling scheme was used to draw samples from the posterior distributions of the parameters. The MCMC run consisted of 20000 cycles, with the first 5,000 cycles discarded as burn-in. The samples of marker effects from the subsequent 15000 cycles were averaged to estimate marker effects. Due to the use of a random model, each scenario was repeated 10 times. The mean and standard deviation of the results were calculated for comparison. Genotypic data quality control was performed using minor allele frequency (MAF). For each replication, markers with MAF less than 0.05 (MAF<0.05) were removed from the genotypic matrix before estimating marker effects.

Model evaluation, cross-validation, predictive accuracy (PA), and bias
The models were evaluated by the goodness of fit, based on the coefficient of determination (R^2), which represents the proportion of variance in the dependent variable predicted by the statistical model. R^2 was calculated as the square of the Pearson's correlation coefficient between the real and estimated values for the full dataset. Additionally, the six gene action models were compared using Akaike's Information Criterion (AIC) (Akaike, 1974) and Deviance Information Criterion (DIC) (Spiegelhalter *et al.* 2002) for each trait.

To divide the dataset into reference and target populations, a 5-fold cross-validation technique (Bengio and Grandvalet, 2004) was used. In each 5-fold repetition, 4 folds were used as the reference population (80% of individuals), and the remaining subset formed the target population (20% of individuals). The phenotypes of the reference dataset were used to derive the genomic predictions (Weber *et al.* 2012), while the phenotypes of the target dataset were ignored. Only the target genotypes were used to derive the genomic predictions. The target group was rotated through all groups until each group was used as the target at least once. Estimations were performed within each fold and averaged across folds and replicates.

The PA was defined as the Pearson's correlation coefficient between the true total genetic values (TGV) or true additive values (true breeding value - TBV) and the genomic estimated total genetic values (GEGV) or predicted

additive values (genomic estimated breeding value - GEBV), i.e., r(TGV, GEGV) or r(TBV, GEBV). For models predicting only additive genetic effects, GEBV estimations were used. For AD and ADE models, PA was calculated separately for both GEBV and GEGV. The simulated TBVs and TGVs are specified in the simulation model and are therefore known exactly.

Bias was calculated as the regression coefficient of TGV on GEGV (TBV on GEBV for model A, and TGV on GEGV for AD and ADE models). The regression coefficient is expected to be close to 1, indicating that the predicted value is similar to the true value. A regression coefficient greater than 1 suggests inflation of genomic predictions, while a value lower than 1 indicates deflation.

The PAs of GEGV (which include additive effects (GEBV), additive effects + dominance deviation, and additive effects + dominance deviation + epistasis interaction for GBLUP or BLASSO, GBLUP-D or BLASSO-D, and GBLUP-DE or BLASSO-DE models, respectively) across 10 repetitions were compared using a t-test for statistical significance at a 0.05 significance level. Practical significance was assessed using Cohen's d effect size. Cohen's d statistic (Cohen, 1988) can be used to describe the standardized mean difference of an effect or to compare effects across studies. It is independent of different ways of measuring the dependent variable. The Cohen's d statistic, which represents the standardized difference between two groups of independent observations, is calculated as follows:

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{S_p}$$

Where:
$\bar{x}_1$ and $\bar{x}_2$: mean PAs of the two compared models.
S_P: pooled standard deviation, calculated as:

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}}$$

Where:
S_1^2 and S_2^2: variances of the PA of the first and second statistical methods, respectively.
n_1 and n_2: number of repetitions of PA in the first and second statistical methods, respectively.

Cohen's d values are interpreted as follows (Cohen, 1988): **0.2**: Small effect, **0.5**: Medium effect and **0.8** or greater: Large effect. A small effect size indicates a weak relationship or difference, while a large effect size indicates a strong relationship or difference implying a practical significance. It ranges from 0 to infinity. By using Cohen's d along with the t-test, both statistical and practical significance were assessed for the different models.

RESULTS AND DISCUSSION

Generally, the lowest and highest AICs or DICs were observed for traits T2 and T5, respectively (Table 2). For trait T1, the lowest AIC and DIC values were obtained for the BLASSO and GBLUP-DE models, respectively. For traits T2 and T4, the lowest AIC and DIC were observed for the GBLUP-DE model. The evaluation of models based on AIC and DIC for traits T3 and T5 showed results similar to those for trait T1.

For trait T6, the lowest AIC and DIC values were similar to those observed for T2 and T4. The models' evaluation based on the coefficient of determination revealed that the GBLUP-DE and BLASSO-DE models outperformed the other gene action models across all traits (Table 3). An increase in heritability led to an improvement in the goodness of fit, while an increase in the number of QTLs resulted in a decline in the goodness of fit.

The PAs of the six studied models, including GBLUP, BLASSO, GBLUP-D, BLASSO-D, GBLUP-DE, and BLASSO-DE, for different traits are presented in Table 4. The PA of all models for all traits increased with increasing heritability and decreased with the increasing number of QTLs.

Based on the obtained results, BLASSO recorded a higher PA than GBLUP. This increase in PA was greater when the number of QTLs was low (20 QTLs), and this advantage decreased as the number of QTLs increased (200 QTLs).

For T1 and T2 traits, the PA of the GBLUP and BLASSO models was slightly higher than that of the other models. In other words, the PA decreased when non-additive effects were included in the model. As the number of QTLs controlling the trait increased, the inclusion of dominance deviation (i.e., GBLUP-D and BLASSO-D) in the model led to an increase in PA. This increase in PA was maximized when epistasis interaction (i.e., GBLUP-DE and BLASSO-DE) was added.

For T1 and T2 traits, no significant difference was observed between the PA of different models (P>0.05) (Table 5). For trait T1, the maximum (d=0.094) and minimum (d=0.818) similarity in PA were observed between GBLUP-D and BLASSO-D, and BLASSO and BLASSO-D models, respectively. As heritability increased (trait T2), the similarity between different models increased. The highest similarity (d=0.014) was observed between BLASSO and BLASSO-DE, while the lowest similarity (d=0.784) was observed between GBLUP-D and BLASSO models for trait T2.

Table 2 Akaike information criterion (AIC) and deviance information criterion (DIC) for each model and trait

Trait	Model	AIC	DIC
T1	GBLUP	1865	1689
	GBLUP-D	1916	1671
	GBLUP-DE	1934	1482
	BLASSO	1856	1691
	BLASSO-D	1871	1682
	BLASSO-DE	1902	1679
T2	GBLUP	1543	1236
	GBLUP-D	1536	1114
	GBLUP-DE	1418	836
	BLASSO	1521	1214
	BLASSO-D	1513	1137
	BLASSO-DE	1517	1108
T3	GBLUP	2553	2404
	GBLUP-D	2603	2376
	GBLUP-DE	2620	2172
	BLASSO	2541	2406
	BLASSO-D	2566	2392
	BLASSO-DE	2582	2388
T4	GBLUP	2158	1890
	GBLUP-D	2105	1665
	GBLUP-DE	1975	1385
	BLASSO	2152	1888
	BLASSO-D	2105	1688
	BLASSO-DE	2118	1715
T5	GBLUP	2987	2843
	GBLUP-D	3033	2804
	GBLUP-DE	3055	2618
	BLASSO	2975	2847
	BLASSO-D	3002	2820
	BLASSO-DE	3010	2805
T6	GBLUP	2438	2182
	GBLUP-D	2399	1979
	GBLUP-DE	2243	1656
	BLASSO	2436	2181
	BLASSO-D	2400	1998
	BLASSO-DE	2398	1995

T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.
GBLUP-D: GBLUP method used in the additive plus dominance deviation model; GBLUP-DE: GBLUP method used in the additive plus dominance deviation model plus epistasis interaction; BLASSO-D: Bayesian LASSO method used in the additive plus dominance deviation model and BLASSO-DE: Bayesian LASSO method used in the additive plus dominance deviation plus epistasis interaction model.

The most significant differences (P<0.05) between different models were observed for T3 and T4 traits, with 12 significant comparisons out of 15 possible comparisons for each trait. The highest similarity in PA for T3 and T4 traits was observed between GBLUP and BLASSO (d=0.112), GBLUP-D and BLASSO-D (d=0.112), and GBLUP and BLASSO (d=0.048), respectively.

In general, for traits controlled by a large number of QTLs, considering non-additive effects in the model led to an increase in the PA of GEBV and GEGV and reduced the bias. The highest and lowest differences in PA were observed between GBLUP and GBLUP-DE (d=5.703, P<0.001) and GBLUP-DE and BLASSO-DE (d=0.003, P>0.05) for T5 and T6 traits, respectively.

The results of the regression coefficients, which indicate the degree of bias in the model, are presented in Table 6. In all traits and different models of gene action, BLASSO showed slightly less bias than GBLUP. For most of the models and traits, the regression coefficient was less than one, indicating the deflation of the GEBV for the GBLUP and BLASSO models and the GEGV in the models that included non-additive effects. Only the GBLUP-DE and BLASSO-DE models for T2 and T4 traits, BLASSO-DE for T3, and GBLUP-DE for T6 inflated the GEGVs.

Table 3 Goodness of fit: The coefficient of determination (R^2) which is calculated as the square of the Pearson's correlation coefficient between the true total genetic values (G_{full}) and the predicted total genetic value ($\hat{G}_{full}$) of the full dataset

Trait	Genetic effects	GBLUP $(r(G_{full}, \hat{G}_{full}))^2$	Bayesian LASSO $(r(G_{full}, \hat{G}_{full}))^2$
T1	Purely add	0.508	0.516
	Add + dom	0.504	0.504
	Add + dom + epis	0.520	0.579
T2	Purely add	0.717	0.726
	add + dom	0.774	0.778
	add + dom + epis	0.839	0.839
T3	Purely add	0.287	0.292
	Add + dom	0.430	0.432
	Add + dom + epis	0.518	0.584
T4	Purely add	0.404	0.406
	Add + dom	0.667	0.669
	Add + dom + epis	0.824	0.832
T5	Purely add	0.112	0.112
	Add + dom	0.161	0.162
	Add + dom + epis	0.494	0.543
T6	Purely add	0.250	0.250
	Add + dom	0.333	0.333
	Add + dom + epis	0.812	0.814

T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.
$(r(G_{full}, \hat{G}_{full}))^2$: the square of the Pearson's correlation coefficient between the true total genetic values (G_{full}) and the predicted total genetic value ($\hat{G}_{full}$) of the full dataset; Purely add: purely additive model (model A); Add + dom: additive plus dominance deviation model (model AD) and Add + dom + epis: additive plus dominance deviation plus epistasis interaction model (model ADE).

Table 4 Predictive accuracy (PA) estimated for genomic estimated breeding values (GEBVs) and genomic estimated genetic values (GEGVs) of the validation population for all models and traits

Trait	Genetic effects	GBLUP r(TBV, GEBV)	GBLUP r(TGV, GEGV)	BLASSO r(TBV, GEBV)	BLASSO r(TGV, GEGV)
T1	Purely add	0.642 (0.03)		0.642 (0.03)	
	Add + dom	0.642 (0.03)	0.621 (0.03)	0.648 (0.03)	0.624 (0.03)
	Add + dom + epis	0.636 (0.03)	0.634 (0.03)	0.646 (0.03)	0.638 (0.04)
T2	Purely add	0.767 (0.03)		0.781 (0.03)	
	add + dom	0.769 (0.03)	0.758 (0.03)	0.782 (0.03)	0.769 (0.03)
	add + dom + epis	0.765 (0.02)	0.764 (0.03)	0.785 (0.02)	0.780 (0.03)
T3	Purely add	0.467 (0.05)		0.472 (0.05)	
	Add + dom	0.500 (0.05)	0.537 (0.03)	0.495 (0.05)	0.541 (0.04)
	Add + dom + epis	0.510 (0.05)	0.602 (0.04)	0.510 (0.05)	0.614 (0.04)
T4	Purely add	0.572 (0.06)		0.575 (0.06)	
	Add + dom	0.617 (0.05)	0.698 (0.04)	0.618 (0.05)	0.702 (0.04)
	Add + dom + epis	0.618 (0.05)	0.745 (0.03)	0.628 (0.05)	0.751 (0.03)
T5	Purely add	0.294 (0.06)		0.297 (0.06)	
	Add + dom	0.326 (0.06)	0.350 (0.08)	0.320 (0.06)	0.354 (0.07)
	Add + dom + epis	0.317 (0.07)	0.552 (0.03)	0.336 (0.06)	0.547 (0.03)
T6	Purely add	0.443 (0.05)		0.444 (0.05)	
	Add + dom	0.482 (0.06)	0.493 (0.07)	0.481 (0.06)	0.495 (0.07)
	Add + dom + epis	0.483 (0.06)	0.677 (0.03)	0.485 (0.06)	0.677(0.04)

T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.
r(TBV, GEBV): pearson's correlation coefficient between the true breeding value (TBV) and the genomic estimated breeding value (GEBV); r(TGV, GEGV): pearson's correlation coefficient between the true genetic values (TGV) and the genomic estimated genetic values (GEGV); Purely add: purely additive model (model A); Add + dom: additive plus dominance deviation model (model AD) and Add + dom + epis: additive plus dominance deviation plus epistasis interaction model (model ADE).

Table 5 Cohen's d effect size output (t-test output (P-value)) between the prediction accuracy of GEGV across 10 replicates for all GBLUP and BLASSO models and traits

Model	T1					T2				
	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE
GBLUP	0.654 (0.161)	0.256 (0.582)	0.191 (0.669)	0.620 (0.183)	0.109 (0.811)	0.318 (0.493)	0.085 (0.852)	0.494 (0.281)	0.079 (0.859)	0.423 (0.361)
GBLUP-D		0.375 (0.41)	0.834 (0.082)	0.094 (0.844)	0.475 (0.298)		0.205 (0.648)	0.784 (0102)	0.390 (0.388)	0.689 (0.141)
GBLUP-DE			0.434 (0.341)	0.315 (0.491)	0.125 (0.751)			0.523 (0.262)	0.154 (0.731)	0.461 (0.323)
BLASS0				0.818 (0.079)	0.276 (0.542)				0.410 (0.371)	0.0141 (0.969)
BLASSO-D					0.425 (0.352)					0.350 (0.441)
	T3					T4				
Model	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE
GBLUP	1.75 (P<0.01)	2.96 (P<0.001)	0.112 (0.801)	1.69 (P<0.01)	3.40 (P<0.001)	2.52 (P<0.001)	3.72 (P<0.001)	0.048 (0.921)	2.61 (P<0.001)	3.86 (P<0.001)
GBLUP-D		1.88 (P<0.001)	1.60 (P<0.01)	0.112 (0.799)	2.46 (P<0.001)		1.42 (P<0.01)	2.47 (P<0.001)	0.112 (0.798)	1.61 (P<0.01)
GBLUP-DE			2.82 (P<0.001)	1.59 (P<0.01)	0.311 (0.302)			3.67 (P<0.001)	1.29 (P<0.01)	0.251 (0.591)
BLASS0				1.55 (P<0.01)	3.26 (P<0.001)				2.56 (P<0.001)	3.81 (P<0.001)
BLASSO-D					2.05 (P<0.001)					1.49 (P<0.01)
	T5					T6				
Model	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE	GBLUP-D	GBLUP-DE	BLASSO	BLASSO-D	BLASSO-DE
GBLUP	0.790 (0.0911)	5.70 (P<0.001)	0.0501 (0.91)	0.892 (0.0611)	5.42 (P<0.001)	0.815 (0.091)	5.28 (P<0.001)	0.023 (0.961)	0.883 (0.0611)	5.22 (P<0.001)
GBLUP-D		4.74 (P<0.001)	0.741 (0.112)	0.0461 (0.921)	3.16 (P<0.001)		3.38 (P<0.001)	0.792 (0.091)	0.041 (0.932)	3.36 (P<0.001)
GBLUP-DE			5.48 (P<0.001)	3.51 (P<0.001)	0.172 (0.712)			5.21 (P<0.001)	3.44 (P<0.001)	0.0031 (0.991)
BLASS0				0.841 (0.123)	5.21 (P<0.001)				0.860 (0.07)	5.16 (P<0.001)
BLASSO-D					3.36 (P<0.001)					3.41 (P<0.001)

T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.
GBLUP-D: GBLUP method used in the additive plus dominance deviation model; GBLUP-DE: GBLUP method used in the additive plus dominance deviation model plus epistasis interaction; BLASSO-D: bayesian LASSO method used in the additive plus dominance deviation model and BLASSO-DE: bayesian LASSO method used in the additive plus dominance deviation plus epistasis interaction model.

All models showed a lower regression coefficient for GEBV than for GEGV across all traits, indicating that when the genetic effects affecting the phenotypes included additive effects, dominance deviations, and epistasis interactions, the bias decreased.

Our main objective in this study was to compare the performance of GBLUP and BLASSO in predicting complex phenotypes under different gene action models and genetic architectures with varying heritability and the number of QTLs. Based on the results of the present study, an increase in heritability and the number of QTLs, respectively, increased and decreased the genomic PA. In a study, it was reported that heritability has the greatest effect on the PA of genomic evaluation, and increasing heritability from 0.15 to 0.45 leads to a significant increase in PA (Atefi *et al.* 2016). In another study, it was suggested that with an increase in heritability from 0.1 to 0.9, the PA of the genome increased from 0.3 to 0.7 (Hayes *et al.* 2010). In a simulated study, it was shown that increasing heritability from 0.2 to 0.6 resulted in a significant increase in PA (Sahebalam *et al.* 2024b).

In a simulation study, it was reported that with an increase in heritability from 0.1 to 0.5, the PA of genomic evaluations increased from 0.52 to 0.77 using the GBLUP method (Sahebalam *et al.* 2022). It has also been reported that by increasing heritability from 0.25 to 1, the PA, in terms of the genetic architecture of the trait, increased from 0.05 to nearly 1 (Combs and Bernardo, 2013).

Table 6 Regression coefficient (b) estimated with total genetic values of the validation population (TGV on GEGV) for all models and traits

Trait	Genetic effects	GBLUP b(TGV, GEGV)	Bayesian LASSO b(TGV, GEGV)
T1	Purely add	0.702 (0.06)	0.721 (0.07)
	Add + dom	0.831 (0.08)	0.859 (0.07)
	Add + dom + epis	0.929 (0.10)	0.948 (0.12)
T2	Purely add	0.809 (0.06)	0.811 (0.06)
	add + dom	0.972 (0.08)	0.982 (0.08)
	add + dom + epis	1.04 (0.07)	1.02 (0.09)
T3	Purely add	0.421 (0.06)	0.441 (0.07)
	Add + dom	0.618 (0.07)	0.689 (0.10)
	Add + dom + epis	1.00 (0.11)	1.05 (0.14)
T4	Purely add	0.491 (0.06)	0.501 (0.05)
	Add + dom	0.772 (0.06)	0.782 (0.06)
	Add + dom + epis	1.06 (0.06)	1.03 (0.06)
T5	Purely add	0.209 (0.06)	0.228 (0.06)
	Add + dom	0.321 (0.09)	0.359 (0.09)
	Add + dom + epis	0.989 (0.08)	0.971 (0.09)
T6	Purely add	0.342 (0.05)	0.342 (0.05)
	Add + dom	0.481 (0.08)	0.488 (0.08)
	Add + dom + epis	1.02 (0.07)	0.971 (0.08)

T1: a trait explained by 20 QTLs with a broad-sense heritability of 0.4; T2: a trait explained by 20 QTLs with a broad-sense heritability of 0.8; T3: a trait explained by 100 QTLs with a broad-sense heritability of 0.4; T4: a trait explained by 100 QTLs with a broad-sense heritability of 0.8; T5: a trait explained by 200 QTLs with a broad-sense heritability of 0.4 and T6: a trait explained by 200 QTLs with a broad-sense heritability of 0.8.
Purely add: Purely additive model (model A); Add + dom: additive plus dominance deviation model (model AD); Add + dom + epis: additive plus dominance deviation plus epistasis interaction model (model ADE) and b(TGV, GEGV): the regression coefficient of TGV on GEGV (TBV on GEBV for model A, and TGV on GEGV for AD and ADE models)

This can be explained by greater genetic variation and less environmental influence in higher heritability, contributing to more PA of marker effects (Barbosa *et al.* 2021).

When the number of QTLs increases, the total genetic variation is expected to be divided among the QTLs, which can reduce the efficiency of methods in estimating small QTL effects and lead to a loss of precision (Resende *et al.* 2012; Ghafouri-Kesbi *et al.* 2016). This is only confirmed for traits that exhibit stronger interactions within the same linkage group, such as traits with 20 QTLs. Since these traits have fewer QTLs in a single linkage group, the expression of interactions between these QTLs is stronger. On the other hand, the decrease in efficiency for a larger number of QTLs can be attributed to the excess of markers with null effects, which can impair the accuracy of the methods (Barbosa *et al.* 2021; Sousa *et al.* 2021). In addition, using simulated data, the researchers showed that the PA of GEBV decreased as the number of QTLs increased from 50 to 200 (Sahebalam *et al.* 2019). In another study, an increase in the number of QTLs led to a decrease in PA (Daetwyler *et al.* 2010).

When the traits were controlled by a small number of QTLs (T1 and T2), the BLASSO method performed better than the GBLUP method. In a simulated study, the PA of GBLUP and BayesC methods was evaluated under different architectures such as the number of QTLs and the distribution of QTL effects. They reported that both methods provided acceptable PA.

However, when the trait phenotype was explained by a small number of QTLs with gamma distribution, the BayesC method performed better (Shirali *et al.* 2015). In another study, using simulated traits with 400 QTL randomly distributed on a genome with 4 chromosomes, the PA and bias of frequentist methods including Ridge Regression, LASSO, Elastic Network and GBLUP, and Bayesian methods including BRR, BayesA, BayesB, BayesCπ and BLASSO were compared. They reported that the Bayesian method had the highest PA and the LASSO and Elastic Net methods had the lowest accuracy. Ridge Regression also recorded the lowest bias (Sahebalam *et al.* 2024a). The lower accuracy of BLUP compared to Bayesian methods for traits with low QTLs could be due to the fact that BLUP uses an infinitesimal model, so all predictors (markers) have normal distribution and partial effects. However, Bayesian methods use different distributions such as t (BayesA), mixture (a point of mass at zero and a Gaussian slab (BayesC) or a point of mass at zero and a scaled-t slab (BayesB)), and double exponential (BLASSO) for markers, which, by not considering all markers equally, have a greater ability to identify effective QTLs, and consequently, have a higher PA in analyzing traits explained by a small number of QTLs.

The results showed that when traits were controlled by a small number of QTLs (20 QTLs), including dominance deviation and epistasis interaction decreased the PA of GEGV compared to GEBV in both studied methods. In a

study using simulated and real data, it was reported that as the contribution of dominance variance increased in total variance, the PA decreased in all studied methods (de Almeida Filho *et al.* 2019). In another study, using a simulated trait with 20 QTLs, it was shown that dominance and epistasis gene activity reduced the PA of genomic evaluation methods. They also reported that the low PAs under non-additive effects could be due to the fact that dominance and epistatic variance are nested within the additive variance. Therefore, the contribution of additive variance to total genetic variance decreases while error variance remains constant (Salehi *et al.* 2021). In addition, researchers reported that the PAs of both parametric and non-parametric methods decreased when gene action was more complex (Momen *et al.* 2018). The decrease in PA due to an increase in the number of QTLs may have occurred due to the greater influence of the multiplicative effect between additive and dominant effects, which is a characteristic of epistatic effects in more complex traits (Coster *et al.* 2010; Barbosa *et al.* 2021). When the number of QTLs was 100, the PA of GEGV increased compared to GEBV, and this difference in accuracy became more evident when the number of QTLs increased to 200. That is, in traits controlled by a large number of QTLs, the inclusion of dominance and epistasis effects improved the PA of both GEBV and GEGV. In a study, reported that a statistical model combining additive and epistatic effects performs better than a simple additive model for the analysis of quantitative traits with epistatic architecture (Morgante *et al.* 2018). In addition, researchers showed that the performance of the GBLUP model, including genetic effects of dominance (GBLUP-D), was evaluated by variance estimation and prediction of genetic merit in computer simulations and two real traits in pigs. In the simulation data, the GBLUP-D model explained more than 50% of the genetic variance of dominance. Additionally, the GBLUP-D model yielded estimated total genetic effects that were 1.2% more accurate than those obtained from GBLUP (Nishio and Satoh, 2014). Also, in another study, it was suggested that for some traits, a larger proportion of the phenotypic variance was explained by non-additive effects compared to additive effects, indicating that epistasis, dominance, or a combination of these effects are of great importance. The genetic effects of epistasis have a greater contribution to the total phenotypic variance than the genetic effects of dominance. Models with non-additive genetic effects did not show obvious superiority over the additive model based on Akaike's information criterion (AIC). Variance component partitioning resulted in the re-ranking of cows relative to the purely additive genetic effects model, indicating that correction for non-additive genetic effects can be effective for selection decisions in dairy cattle breeding programs. These results showed that non-additive genetic effects play an important role in some fertility and reproductive traits in Holstein dairy cows (Alves *et al.* 2020).

In the present study, the regression coefficient values for most of the models and traits were less than one, and only for additive plus dominance plus epistasis gene action were they greater than one in some traits. In a study on several egg production traits, it was reported that all regression coefficient values for both GBLUP and BayesC methods were less than one (Heidaritabar *et al.* 2016). In another study on a population of Holstein dairy cattle, the regression coefficient was higher than one for the GBLUP method (Charfeddine *et al.* 2013). In addition, researchers showed that the regression coefficient was lower than one for the GBLUP method on French Lacunae dairy sheep (Duchemin *et al.* 2012).

The main limitations of this study include the reliance on simulated data, which may not fully represent the complexity of real-world breeding programs. Additionally, the study primarily evaluated prediction accuracy and bias, without considering other important metrics such as model stability or performance in real-world conditions with environmental interactions. The computational burden of non-additive models like GBLUP-DE and BLASSO-DE also raises concerns for scaling up to larger datasets. Lastly, while the models performed well in controlled settings, their robustness in real-world applications, with issues like incomplete data and marker biases, requires further investigation. These limitations suggest the need for future research using real-world datasets and exploring the practical applicability of these models.

CONCLUSION

Our study demonstrates that both GBLUP and BLASSO methods exhibit similar performance in predicting complex traits, with only slight differences in PA and bias. Heritability and the number of QTLs significantly influenced the PA of GEBV and GEGV. Specifically, higher heritability increased PA, while a greater number of QTLs generally reduced PA. For traits controlled by a small number of QTLs, the purely additive model was sufficient to achieve acceptable PA. However, when traits involved a larger number of QTLs, incorporating non-additive effects, such as dominance and epistasis interactions, enhanced both PA and reduced bias. These findings underscore the importance of considering gene action models in genomic evaluations, particularly for traits with complex genetic architectures.

ACKNOWLEDGEMENT

We are grateful to the anonymous reviewers for their comments which improved the quality of the manuscript.

REFERENCES

Abdollahi-Arpanahi R., Gianola D. and Peñagaricano F. (2020). Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. *Genet. Sel. Evol*. **52,** 1-15.

Akaike H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Cont*. **19(6),** 716-723.

Aliloo H., Pryce J.E., González-Recio O., Cocks B.G. and Hayes B.J. (2016). Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. *Genet. Sel. Evol*. **48,** 1-11.

Aliloo H., Pryce J.E., González-Recio O., Cocks B.G., Goddard M.E. and Hayes B.J. (2017). Including nonadditive genetic effects in mating programs to maximize dairy farm profitability. *J. Dairy. Sci*. **100,** 1203-1222.

Alves K., Brito L.F., Baes C.F., Sargolzaei M., Robinson J.A.B. and Schenkel F.S. (2020). Estimation of additive and non-additive genetic effects for fertility and reproduction traits in North American Holstein cattle using genomic information. *J. Anim. Breed. Genet*. **137,** 316-330.

Atefi A., Shadparvar A.A. and Hossein-Zadeh N.G. (2016). Comparison of whole genome prediction accuracy across generations using parametric and semi parametric methods. *Acta Sci. Anim. Sci.* **38,** 447-453.

Barbosa I.D.P., da Silva M.J., da Costa W.G., de Castro Sant'Anna I., Nascimento M. and Cruz C.D. (2021). Genome-enabled prediction through machine learning methods considering different levels of trait complexity. *Crop Sci*. **61,** 1890-1902.

Bengio Y. and Grandvalet Y. (2004). No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res*. **5,** 1089-1105.

Charfeddine N., Rodríguez-Ramilo S.T., Jiménez-Montero J.A., Carabaño M.J. and González-Recio O. (2013). Non parametric vs. gblup model for genomic evaluation with large reference population in holstein cattle. *Int. Bull*. **47,** 25-35.

Cohen J. (1988). Statistical Power Analysis for the Behavioral Sciences. Routledge Academic, New York.

Combs E. and Bernardo R. (2013). Accuracy of genome wide selection for different traits with constant population size, heritability, and number of markers. *Plant Genome*. **6,** 1-7.

Coster A., Bastiaansen J.W., Calus M.P., van Arendonk J.A. and Bovenhuis H. (2010). Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. *Genet. Sel. Evol*. **42,** 1-11.

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

de Almeida Filho J.E., Guimarães J.F.R., Fonsceca e Silva F., Vilela de Resende M.D., Muñoz P., Kirst M. and de Resende Júnior M.F.R. (2019). Genomic prediction of additive and non-additive effects using genetic markers and pedigrees. *G3: Genes, Genomes, Genetics*. 9, 2739-2748.

Duchemin S.I., Colombani C., Legarra A., Baloche G., Larroque H., Astruc J.M., Barillet F., Robert-Granié C. and Manfredi E. (2012). Genomic selection in the French Lacaune dairy sheep breed. *J. Dairy. Sci*. **95**, 2723-2733.

Duenk P., Calus M.P., Wientjes Y.C. and Bijma P. (2017). Benefits of dominance over additive models for the estimation of average effects in the presence of dominance. *G3: Genes, Genomes Genet*. **7,** 3405-3414.

Ghafouri-Kesbi F., Rahimi-Mianji G., Honarvar M. and Nejati-Javaremi A. (2016). Predictive ability of random forests, boosting, support vector machines and genomic best linear unbiased prediction in different scenarios of genomic evaluation. *Anim. Prod. Sci*. **57,** 229-236.

Habier D., Fernando R.L. and Dekkers J. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*. **177,** 2389-2397.

Hayes B.J., Bowman P.J., Chamberlain A.C., Verbyla K. and Goddard M.E. (2009a). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol*. **41,** 1-9.

Hayes B.J., Daetwyler H.D., Bowman P., Moser G., Tier B., Crump R., Khatkar M., Raadsma H.W. and Goddard M.E. (2010). Accuracy of genomic selection: comparing theory and results. *Proc. Assoc. Advmt. Anim. Breed. Genet*. **18,** 34-37.

Hayes B.J., Visscher P.M. and Goddard M.E. (2009b). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* **91,** 47-60.

Heidaritabar M., Wolc A., Arango J., Zeng J., Settar P., Fulton J.E., O'Sullivan N.P., Bastiaansen J.W., Fernando R.L., Garrick D.J. and Dekkers J.C. (2016). Impact of fitting dominance and additive effects on accuracy of genomic prediction of breeding values in layers. *J. Anim. Breed. Genet*. **133,** 334-346.

Hill W.G. (2010). Understanding and using quantitative genetic variation. *Philos. Trans. R. Soc. Lond. B Sci*. **365,** 73-85.

Legarra A., Robert-Granié C., Manfredi E. and Elsen J.M. (2008). Performance of genomic selection in mice. *Genetics*. **180,** 611-618.

Mäki-Tanila A. (2007). An overview on quantitative and genomic tools for utilizing dominance genetic variation in improving animal production. *Agric. Food Sci.* **16,** 188-198.

McHugh N., Meuwissen T.H.E., Cromie A.R. and Sonesson A.K. (2011). Use of female information in dairy cattle genomic breeding programs. *J. Dairy Sci.* **94,** 4109-4118.

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

Momen M., Mehrgardi A.A., Sheikhi A., Kranis A., Tusell L.,

Morota G., Rosa G.J. and Gianola D. (2018). Predictive ability of genome-assisted statistical models under various forms of gene action. *Sci. Rep*. **8,** 12309-12318.

Morgante F., Huang W., Maltecca C. and Mackay T.F. (2018). Effect of genetic architecture on the prediction accuracy of quantitative traits in samples of unrelated individuals. *Heredity*. **120,** 500-514.

Nishio M. and Satoh M. (2014). Including dominance effects in the genomic BLUP method for genomic evaluation. *PLoS One*. **9,** e85792.

Park T. and Casella G. (2008). The bayesian lasso. *J. Am. Stat. Assoc*. **103,** 681-686.

Pérez P. and de Los Campos G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*. **198,** 483-495.

Resende M.D., Resende M.F., Sansaloni C.P., Petroli C.D., Missiaggia A.A., Aguiar A.M., Abad J.M., Takahashi E.K., Rosado A.M., Faria D.A. and Pappas G.J. (2012). Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol*. **194,** 116-128.

Sahebalam H., Gholizadeh M., Hafezian H. and Farhadi A. (2019). Comparison of parametric, semiparametric and non-parametric methods in genomic evaluation. *J. Genet*. **98,** 1-8.

Sahebalam H., Gholizadeh M., Hafezian H. and Ebrahimi F. (2022). Evaluation of Bagging approach versus GBLUP and Bayesian LASSO in genomic prediction. *J. Genet*. **101,** 19-28.

Sahebalam H., Gholizadeh M. and Hafezian H. (2024a). Investigating the performance of frequentist and Bayesian techniques in genomic evaluation. *Biochem. Genet*. **21,** 1-27.

Sahebalam H., Gholizadeh M. and Hafezian S.H. (2024b). The effect of different approaches to determining the regularization parameter of bayesian LASSO on the accuracy of genomic prediction. *Mamm. Genome*. **18,** 1-15.

Salehi A., Bazrafshan M. and Abdollahi-Arpanahi R. (2021). Assessment of parametric and non-parametric methods for prediction of quantitative traits with non-additive genetic architecture. *Ann. Anim. Sci.* **21**, 469-484.

Shirali M., Miraei-Ashtiani S.R., Pakdel A., Haley C., Navarro P. and Pong-Wong R. (2015). A comparison of the sensitivity of the BayesC and genomic best linear unbiased prediction (GBLUP) methods of estimating genomic breeding values under different quantitative trait locus (QTL) model assumptions. *Iranian J. Appl. Anim. Sci*. **5,** 41-46.

Sousa I.C.D., Nascimento M., Silva G.N., Nascimento A.C.C., Cruz C.D., Almeida D.P.D., Pestana K.N., Azevedo C.F., Zambolim L. and Caixeta E.T. (2021). Genomic prediction of leaf rust resistance to Arabica coffee using machine learning algorithms. *Sci. Agric*. **78,** 1-8.

Spiegelhalter D.J., Best N.G., Carlin B.P. and Van der Linde A. (2002). Bayesian measures of model complexity and fit. *J. R. Statist. Soc. B.* **64,** 583-639.

Su G., Christensen O.F., Ostersen T., Henryon M. and Lund M.S. (2012). Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS One*. **7**, e45293.

Toro M.A. and Varona L. (2010). A note on mate allocation for dominance handling in genomic selection. *Genet. Sel. Evol*. **42,** 1-9.

Valdar W., Solberg L.C., Gauguier D., Burnett S., Klenerman P., Cookson W.O., Taylor M.S., Rawlins J.N.P., Mott R. and Flint J. (2006a). Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet*. **38,** 879-887.

Valdar W., Solberg L.C., Gauguier D., Cookson W.O., Rawlins J.N.P., Mott R. and Flint J. (2006b). Genetic and environmental effects on complex traits in mice. *Genetics*. **174,** 959-984.

VanRaden P.M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci*. **91,** 4414-4423.

VanRaden P.M., Van Tassell C.P., Wiggans G.R., Sonstegard T.S., Schnabel R.D., Taylor J.F. and Schenkel F.S. (2009). Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci*. **92,** 16-24.

Varona L., Legarra A., Toro M.A. and Vitezica Z.G. (2018). Non-additive effects in genomic selection. *Front. Genet*. **9,** 78-87.

Vitezica Z.G., Varona L. and Legarra A. (2013). On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics*. **195,** 1223-1230.

Vitezica Z.G., Reverter A., Herring W. and Legarra A. (2018). Dominance and epistatic genetic variances for litter size in pigs using genomic models. *Genet. Sel. Evol*. **50,** 1-8.

Weber K.L., Thallman R.M., Keele J.W., Snelling W.M., Bennett G.L., Smith T.P.L., McDaneld T.G., Allan M.F., Van Eenennaam A.L. and Kuehn L. (2012). Accuracy of genomic breeding values in multibreed beef cattle populations derived from deregressed breeding values and phenotypes. *J. Anim. Sci*. **90,** 4177-4190.

Wiggans G.R., VanRaden P.M. and Cooper T.A. (2011). The genomic evaluation system in the United States: Past, present, future. *J. Dairy Sci*. **94,** 3202-3211.

Zeng J., Toosi A., Fernando R.L., Dekkers J.C. and Garrick D.J. (2013). Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genet. Sel. Evol*. **45,** 1-17.