

Performance of four genomic selection methods using different heritability and QTL numbers

Abstract – The objective of this work was to evaluate how heritability and the number of quantitative trait loci (QTL) controlling the trait can influence the prediction of genetic value by genomic selection methods. A prediction equation was established to estimate genetic correlation based on phenotypic correlation, using an F_2 population with 1,000 individuals, simulated in different scenarios. Heritability (5, 20, 40, 60, 80, and 99%) and QTL number (60, 120, 180, and 240) varied in each scenario. The following four genomic selection methods were used in the analyses: ridge-regression best linear unbiased prediction (RR-BLUP), genomic BLUP (GBLUP), Bayesian estimation method B (Bayes B), and reproducing kernel Hilbert spaces regression (RKHS). The phenotypic and genotypic predictive abilities were calculated for each method, and Tukey's test was used to compare means. The effect of heritability and of the number of QTL controlling the trait was evaluated by the regression analysis. Tukey's test revealed differences between the methods, with Bayes B and RR-BLUP being superior to the others in almost all scenarios. Heritability presents a positive linear relationship with phenotypic predictive ability and a positive quadratic relationship with genotypic predictive ability. The number of QTL controlling the trait has no relationship with the phenotypic and genotypic predictive abilities.

Index terms: accuracy, genome-wide selection, heritability, mixed model, QTL.

Desempenho de quatro métodos de seleção genômica com uso de diferentes herdabilidades e números de QTL

Resumo – O objetivo deste trabalho foi avaliar como a herdabilidade e o número de locos de características quantitativas (QTL) que controla a característica podem influenciar na predição do valor genético por meio de métodos de seleção genômica. Uma equação de predição foi estabelecida para estimar a correlação genética baseada na correlação fenotípica, tendo-se utilizado uma população F_2 com 1.000 indivíduos, simulados em diferentes cenários. A herdabilidade (5, 20, 40, 60, 80 e 99%) e o número de QTL (60, 120, 180 e 240) variaram em cada cenário. Os quatro seguintes métodos de seleção genômica foram utilizados nas análises: *ridge-regression best linear unbiased prediction* (RR-BLUP), BLUP genômico (GBLUP), método bayesiano de estimação B (Bayes B) e *reproducing kernel Hilbert spaces regression* (RKHS). As habilidades preditivas fenotípicas e genotípicas foram calculadas para cada método, e o teste de Tukey foi utilizado para comparação de médias. O efeito da herdabilidade e do número de QTL que controla a característica foi avaliado por análise de regressão. O teste de Tukey revelou diferenças entre os métodos, sendo que Bayes B e RR-BLUP foram superiores aos demais em quase todos os cenários. A herdabilidade apresenta relação linear positiva

Leonardo de Azevedo Peixoto⁽¹⁾ ,
Paulo Eduardo Teodoro⁽²⁾ ,
Larissa Pereira Ribeiro Teodoro⁽²⁾ ,
Cosme Damião Cruz⁽³⁾  and
Leonardo Lopes Bhering⁽³⁾ 

⁽¹⁾ Iowa State University, 716 Farm House Lane,
50011-1051 Ames, IA, United States.
E-mail: leonardo@iastate.edu

⁽²⁾ Universidade Federal de Mato Grosso do
Sul, Campus de Chapadão do Sul, Rodovia
MS-306, Km 105, Caixa Postal 112,
CEP 79560-000 Chapadão do Sul, MS,
Brazil.
E-mail: eduteodoro@hotmail.com,
larissa.uems@gmail.com

⁽³⁾ Universidade Federal de Viçosa,
Departamento de Biologia Geral,
Avenida Peter Henry Rolfs, s/nº, Campus
Universitário, CEP 36570-900 Viçosa, MG,
Brazil.
E-mail: cdcruz@ufv.br,
leonardo.bhering@ufv.br

✉ Corresponding author

Received
September 30, 2023

Accepted
June 20, 2024

How to cite
PEIXOTO, L. de A.; TEODORO, P.E.;
TEODORO, L.P.R.; CRUZ, C.D.; BHERING,
L.L. Performance of four genomic selection
methods using different heritability and
QTL numbers. *Pesquisa Agropecuária
Brasileira*, v.59, e03535, 2024. DOI: <https://doi.org/10.1590/S1678-3921.pab2024.v59.03535>.

com a capacidade preditiva fenotípica e relação quadrática positiva com a capacidade preditiva genotípica. O número de QTL controlando a característica não tem relação com a capacidade preditiva fenotípica e genotípica.

Termos para indexação: precisão, seleção genômica ampla, herdabilidade, modelo misto, QTL.

Introduction

Until 30 years ago, the selection of superior genotypes in most plant and animal breeding programs was based on the visual selection of individuals (Lichhane et al., 2022). This changed with the advent of molecular markers, which allowed of the incorporation of molecular information to improve prediction and selection accuracy (Lichhane et al., 2022). The first marker-based methodology used in breeding was the molecular marker-assisted selection (Xu & Croux, 2008). However, this methodology was useful only for traits with quantitative trait loci (QTL) of major effect, being inefficient for traits controlled by minor-effect genes (Zhong et al., 2009; Song et al., 2023).

With the evolution and the introduction of molecular markers, such as single nucleotide polymorphisms and diversity arrays technology, new statistical models, known as genomic selection models, were established for the study of the influence of minor-effect genes (Meuwissen et al., 2001). These models use the effect of all markers available to estimate the genomic estimated breeding value (GEBV) of an individual.

The prediction accuracy of these models is influenced by several factors, such as the heritability of the trait and the number of genes controlling it (Ornella et al., 2012; Robert et al., 2022; De Mori & Ciprinai, 2023). According to Zhong et al. (2009) and Zargar et al. (2015), in genomic selection methods, accuracy seems to be inversely related to the number of QTL. For example, when estimated by Bayesian methods, accuracy is higher for traits controlled by fewer major-effect genes. Conversely, in models based on the best linear unbiased prediction (BLUP), a better performance is observed for traits controlled by several minor-effect genes (Meuwissen et al., 2001; Zhong et al., 2009). Although there are studies comparing genomic selection methods (Heslot et al., 2012; Bhering et al., 2015), only a few of them have taken heritability and the number of QTL controlling the trait into account (Desta & Ortiz, 2014), whereas

none of them have considered these two factors simultaneously.

The objective of this work was to evaluate how heritability and the number of QTL controlling the trait can influence the prediction of genetic value by genomic selection methods.

Materials and Methods

For the study, an F_2 population was simulated using the simulation module of the GENES software (Cruz, 2013), which allowed of generating information on the genome, the genotypes of the parents, the controlled cross populations, and quantitative trait data. A genome consisting of 15 linkage groups, similar to that of a $2n = 2x = 30$ diploid species, was simulated. Each linkage group had 200 cM, with 200 markers per linkage group, spaced equally at 1 cM, totaling 3,000 markers. The markers were assumed as codominant and biallelic.

Contrasting homozygote parents were simulated, i.e., parent 1 was coded as dominant (2), and parent 2 was coded as recessive (0) for all markers. Therefore, the cross between parent 1 and parent 2 generated the F_1 population with all genes in heterozygosis. The simulated F_2 population was coded with 0, 1, and 2, where 0 corresponds to recessive homozygote individuals, 1 to heterozygote individuals, and 2 to homozygote individuals for a given locus.

The F_2 population was composed of 1,000 individuals, generated from the cross-selfing of individuals of the F_1 population. In this process, each individual of the F_1 population produced 5,000 gametes, and, when 2 of these gametes met at random, the first individual of the F_2 population was generated. This process was repeated until all individuals of each population were formed.

Traits controlled by different QTL numbers (60, 120, 180, and 240) were simulated to verify how the number of QTL controlling the trait could influence the prediction of genetic value by genomic selection methods.

A binomial distribution was assigned to the importance of each QTL, using the following equation:

$$\text{QTL importance} = [n! / k! (n - k)!] \times p^k \times q^{(n-k)}$$

where $q=0.5$; and $N = n - 1$, where n is the number of QTL. This distribution was adopted since it considers

that there are some more important QTL, but that these are not frequent and do not have major effects. This fact makes the simulation more realistic for the study.

The expression of each QTL was defined by: $AA = \mu + a$; $Aa = \mu + d$; $aa = \mu - a$. Since the value of d was defined as null, the mean degree of dominance (d/a) was zero for all loci.

The genotypic value (GV) of each individual was established by the following equation:

$$GV = \sum_{i=1}^n (QTL_i \text{ importance} \times QTL_i \text{ expression})$$

The environment effect was defined as a vector independent of the genotypic value and was estimated following $N(0, \sigma^2)$, where σ^2 is variance, whose value was calculated from the heritability of the traits and the value of genetic variance (σ_g^2). The heritability value was previously defined. Traits with a heritability of 5, 20, 40, 60, 80, and 99% were simulated in the present work, and σ_g^2 was calculated as the variance of the genotypic value of the individuals in the F_2 population.

The phenotypic value (PV) was calculated as follows: $PV = \mu + GV + EV$, where μ is the mean defined by the user ($\mu = 100$ for the present study), and EV is the environmental value.

The mapping process was carried out after the population was generated, starting with the analysis of segregation of individual loci. Chi-square tests were applied to verify if the markers generated in the study segregated according to an F_2 population. All linkage groups were checked for restoration, considering size, distance, and order of markers, which confirmed that the F_2 population had the desired simulation properties.

For the analyses, the following four genomic selection methods, widely used in plant and animal breeding, were tested: ridge-regression BLUP (RR-BLUP), genomic BLUP (GBLUP), Bayesian estimation method B (Bayes B), and reproducing kernel Hilbert spaces regression (RKHS).

RR-BLUP and Bayes B were described by Meuwissen et al. (2001). RR-BLUP assumes that each marker has a variance equal to $GVar/M$, where GV is genetic variance and M is the number of markers. In the Bayes B method, the priori of the proportion of markers associated with the phenotypic variance equal to zero assumes an inverted chi-square distribution.

In RKHS, the genetic values are estimated by the Gaussian process, and all parameters of the priori are described by De Los Campos et al. (2010).

For the comparison of the genomic selection methods, the phenotypic and genotypic predictive abilities were defined as Pearson's correlation between the phenotypic value and the GEBV and as Pearson's correlation between the true genetic value and the GEBV, respectively. In addition, Tukey's test was used for mean comparisons, at 5% probability, for each used scenario.

The regression analysis (through linear, quadratic, and cubic regression models) was used to verify the influence of heritability and of the number of QTL controlling the trait on the prediction accuracy of the tested genomic selection methods, which were evaluated with different heritability values (5, 20, 40, 60, 80, and 99%) and numbers of QTL simulated (60, 120, 180, and 240).

The linear, quadratic, and cubic regression models were tested to predict the genetic correlation (Pearson's correlation between the true genetic value and the GEBV) from the phenotypic correlation (Pearson's correlation between the phenotypic value and the GEBV).

All analyses were performed using the R statistical software (R Core Team, 2017), as follows: RR-BLUP and GBLUP, with `mixed.solve` and `kin`; BLUP functions in the `rrBLUP` package; and Bayes B and RKHS using the `BGLR` function in the `BGLR` package. A total of 20,000 burn-ins and 100,000 MCMC iterations were used in the Bayesian analysis. The convergence of the Bayesian models was analyzed using the variance parameters of the trace plot.

Results and Discussion

Significant differences were observed between the genomic selection methods for all heritability values evaluated, regardless of the number of QTL for the phenotypic (Table 1) and genotypic (Table 2) predictive abilities. In almost all evaluated scenarios, both the phenotypic and genotypic predictive abilities of GBLUP and RKHS were inferior to those of the other methods, whereas those of the RR-BLUP and Bayes B were significantly superior. For heritability values above 40%, the Bayes B method was superior to RR-BLUP.

According to the literature, the performance of a model is strongly influenced by interallelic interaction. For resistance to wheat rust, Ornella et al. (2012) found

that the Bayesian Lasso and Bayesian ridge regression presented superior results to that of the support vector regression, a non-parametric method as the RKHS used in the present study. The authors concluded that parametric methods, such as RR-BLUP, Bayes B, and GBLUP, are superior because the studied trait is controlled by an additive gene effect. Contrastingly, non-parametric methods, as RKHS, can capture non-additive effects, such as dominance and epistasis, but

may even decrease accuracy when the trait has an additive gene control, as verified by Zhao et al. (2013) and in the present study, where RKHS presented lower results in most of the evaluated scenarios. The fact that all traits were simulated with only the additive effect may have led all used methods to present similar results, except the non-parametric RKHS (Tables 1 and 2). Heslot et al. (2012), working with maize (*Zea mays* L.) and barley (*Hordeum vulgare* L.), compared

Table 1. Estimate of the phenotypic predictive ability with different values of heritability (h^2) and numbers of quantitative trait loci (QTL) of the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods⁽¹⁾.

Method	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
	QTL number = 60				QTL number = 120			
h ² = 5%	0.04a	0.05a	0.05a	0.02b	0.16b	0.16b	0.16b	0.17a
h ² = 20%	0.37a	0.37a	0.35b	0.35b	0.34a	0.35a	0.34b	0.34b
h ² = 40%	0.56b	0.58a	0.55c	0.55c	0.56a	0.57a	0.56a	0.56a
h ² = 60%	0.71b	0.73a	0.71b	0.71b	0.69b	0.70a	0.69b	0.69b
h ² = 80%	0.84b	0.86a	0.83c	0.82c	0.80b	0.85a	0.78c	0.79c
h ² = 99%	0.96b	0.99a	0.95c	0.95c	0.96b	0.99a	0.95c	0.96b
	QTL number = 180				QTL number = 240			
h ² = 5%	0.19a	0.19a	0.18b	0.18b	0.14a	0.14a	0.13b	0.13b
h ² = 20%	0.39a	0.39a	0.38b	0.38b	0.32a	0.32a	0.31b	0.31b
h ² = 40%	0.52a	0.52a	0.51b	0.51b	0.54b	0.55a	0.53c	0.53c
h ² = 60%	0.69b	0.70a	0.69b	0.69b	0.69a	0.69a	0.68b	0.68b
h ² = 80%	0.84b	0.86a	0.84b	0.84b	0.79c	0.86a	0.85b	0.85b
h ² = 99%	0.97b	0.99a	0.96c	0.96c	0.97b	0.99a	0.96c	0.97b

⁽¹⁾Means followed by equal letters, in the lines, do not differ by Tukey's test, at 5% probability.

Table 2. Estimate of the genotypic predictive ability with different values of heritability (h^2) and numbers of quantitative trait loci (QTL) of the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods⁽¹⁾.

Method	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
	QTL number = 60				QTL number = 120			
$h^2 = 5\%$	0.56a	0.55b	0.51d	0.53c	0.64a	0.64a	0.60c	0.63b
$h^2 = 20\%$	0.77b	0.78a	0.76c	0.76c	0.75a	0.75a	0.74b	0.74b
$h^2 = 40\%$	0.89b	0.92a	0.88c	0.88c	0.86a	0.86a	0.85b	0.85b
$h^2 = 60\%$	0.91b	0.93a	0.91c	0.91c	0.91b	0.92a	0.91b	0.91b
$h^2 = 80\%$	0.94b	0.97a	0.93c	0.93c	0.90b	0.96a	0.89c	0.89c
$h^2 = 99\%$	0.97b	0.99a	0.95c	0.96c	0.97b	0.99a	0.96c	0.96c
	QTL number = 180				QTL number = 240			
$h^2 = 5\%$	0.64b	0.64b	0.63c	0.65a	0.56a	0.56a	0.52b	0.52b
$h^2 = 20\%$	0.80a	0.80a	0.79b	0.79b	0.77a	0.77a	0.75b	0.74c
$h^2 = 40\%$	0.85b	0.86a	0.84c	0.84c	0.86b	0.87a	0.85c	0.85c
$h^2 = 60\%$	0.89b	0.91a	0.88c	0.88c	0.90b	0.91a	0.89c	0.89c
$h^2 = 80\%$	0.94b	0.96a	0.94b	0.94b	0.95b	0.96a	0.94c	0.94c
$h^2 = 99\%$	0.97b	0.99a	0.96c	0.96c	0.98b	0.99a	0.97c	0.97c

⁽¹⁾Means followed by equal letters, in the lines, do not differ by Tukey's test, at 5% probability.

11 genomic selection methods, separating them into two groups: parametric and non-parametric. In the present work, the RKHS method was classified in a group different from that of the other methods.

Another factor that made the GBLUP and RR-BLUP traditional methods present similar results to that of the Bayesian method was the use of non-informative priori due to the default of the used BGLR package. When the non-informative priori is used, the posteriori is based only on the likelihood function, i.e., although the method was Bayesian, the results only transcribed the likelihood function in the same way that the traditional methods do (Bhering et al., 2015). Moreover, all Bayesian methods of genomic selection use the same original model, and the only difference between them is the hyperparameters in the priori (Xu et al., 2021). Since the priori was non-informative, the methods ended up being very similar, and, consequently, did not present significant differences.

If the priori information on the trait under study is not available, the RR-BLUP and GBLUP traditional methods can be used in the prediction of genetic value. Otherwise, when the priori information is available, the use of Bayesian methods will present better results (Meuwissen et al., 2001). However, if, in addition to the priori information, dominance and/or epistatic effects are also being estimated, the RKHS method is more appropriate.

The values of the phenotypic and genotypic predictive abilities increased with the increase in the heritability value, regardless of the method used or of the number of QTL controlling the trait (Figures 1 and 2).

For the phenotypic predictive ability, there was a positive linear relationship with heritability in all scenarios with a different number of QTL. In addition, the value of the coefficient of determination (R^2) of the linear regression was higher than 0.94 for all the genomic selection methods tested.

For the genotypic predictive ability, the relationship with heritability was quadratic in all scenarios with a different number of QTL controlling the trait. A plateau was reached when the heritability of the trait reached 60% (Figure 2). The R^2 value of the quadratic regression was higher than 0.93 for all the genomic selection methods used.

The correlation between heritability and accuracy is positive, as verified in wheat for yellow rust and stem

rust (Ornella et al., 2012), as well as in maize for grain yield and grain moisture (Zhao et al., 2013). However, heritability and the number of QTL controlling the trait are correlated factors, and, sometimes, traits with a lower heritability value and a higher QTL number present a higher accuracy than those with a higher heritability and a lower QTL number, as noted by Heffner et al. (2011). Similarly, in the present work, all traits controlled by 240 QTL presented a higher accuracy than those controlled by 60 QTL, regardless of heritability, although this relationship was not linear (Figures 3 and 4). However, for traits controlled by the same QTL number, the higher the heritability value, the higher were the phenotypic (Figure 1) and genotypic (Figure 2) accuracies, representing a linear and a quadratic relationship, respectively.

In breeding programs, selection accuracy can be significantly improved through genomic selection (Voss-Fels et al., 2019), mainly for traits with high phenotypic evaluation costs (protein and oil contents, for example) or that are very complex (resistant to diseases) due to their usually low heritability (lower than 30%), which makes selection based only on phenotype very difficult. Therefore, as observed in the present work, the lower the heritability value, the greater the difference between the reliability (square of the predictive accuracy) and heritability of a trait, i.e., for low heritability traits, selection based on the GEBV predicted by the genomic selection methods will be much more accurate than selection based on phenotypic values.

In the different scenarios simulated by varying the number of QTL for the prediction of genetic value, the R^2 values of the cubic regressions, ranging from 0.58 to 0.97, were higher than those of the other regression models (Figure 3). Therefore, no relationship was observed between the number of QTL controlling the trait and phenotypic predictive ability, regardless of the heritability of the trait.

When a trait is controlled by a low number of QTL of major effect, the Bayesian method showed a better performance than the GBLUP and RR-BLUP traditional methods, whereas the opposite was observed when the number of QTL was high (Meuwissen et al., 2001; Zhong et al., 2009). However, this difference may be more influenced by other traits, such as heritability, training population size, and population structure, rather than by QTL number (Desta & Ortiz, 2014).

The obtained results show that studying a factor separately can super- or underestimate the values estimated by genomic selection methods. Therefore, further works should be carried out considering several factors simultaneously, in order to establish

the best genomic selection model for each population structure, which, in the present study, was F_2 .

For the genotypic predictive ability, no relationship was observed between the values predicted by the genomic selection methods and the number of QTL

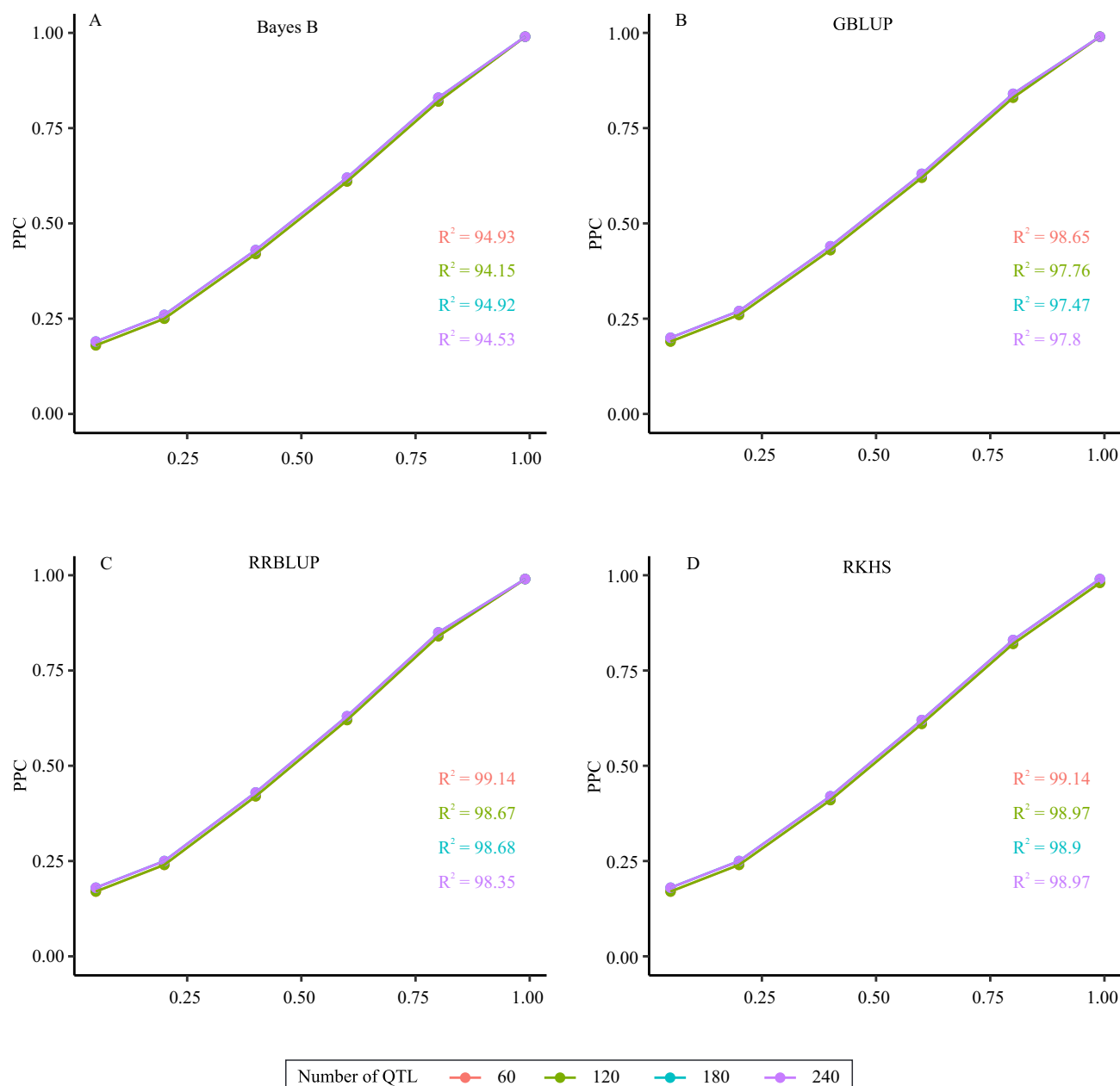


Figure 1. Phenotypic predictive ability (PPC) in function of heritability, with different numbers of QTL controlling the trait, of the following four genomic selection methods: Bayesian estimation method B (A), genomic best linear unbiased prediction (B), reproducing kernel Hilbert spaces regression (C), and ridge-regression best linear unbiased prediction (D). R^2 , coefficient of determination.

controlling the quantitative trait. Once again, cubic regression presented the best results, with a R^2 ranging from 85.29 to 98.58% (Figure 4).

Between the genetic and phenotypic correlations (Tables 3, 4, 5, and 6), a low R^2 value was verified

for the linear, quadratic, and cubic regression models evaluated, regardless of the genomic selection method used to estimate both correlations. The exception was the 99% heritability, which resulted in a R^2 value higher than 89% for RR-BLUP, RKHS, and GBLUP.

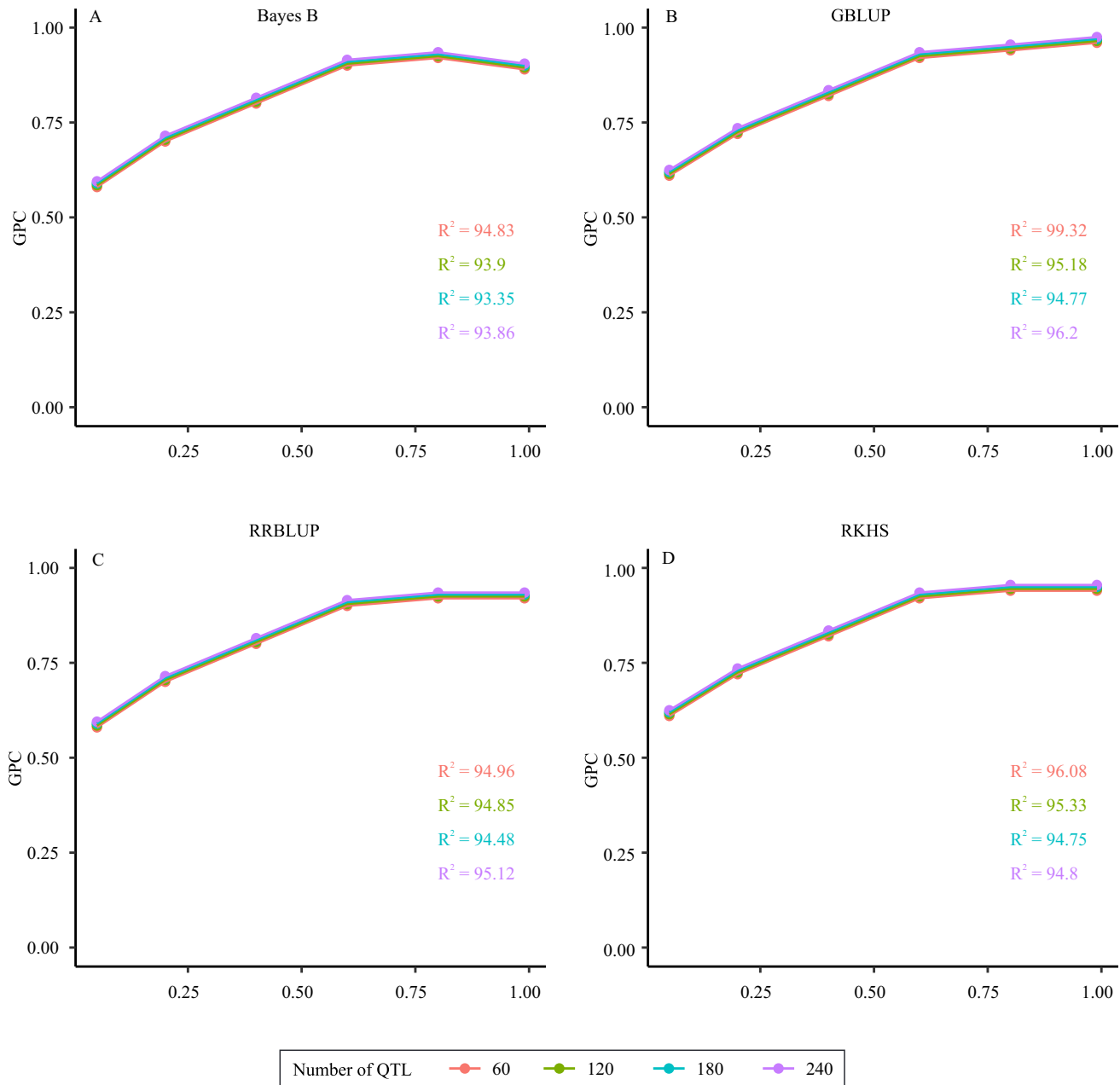


Figure 2. Genotypic predictive ability (GPC) in function of heritability, with different numbers of QTL controlling the trait, of the following four genomic selection methods: Bayesian estimation method B (A), genomic best linear unbiased prediction (B), reproducing kernel Hilbert spaces regression (C), and ridge-regression best linear unbiased prediction (D). R^2 , coefficient of determination.

Moreover, the genomic selection methods differed regarding the prediction of genetic correlation by phenotypic correlation. However, no pattern was detected between the methods, with the results

of R^2 being completely random. The R^2 values of the regressions increased with the increase in the heritability value for almost all evaluated scenarios with varying numbers of QTL controlling the trait,

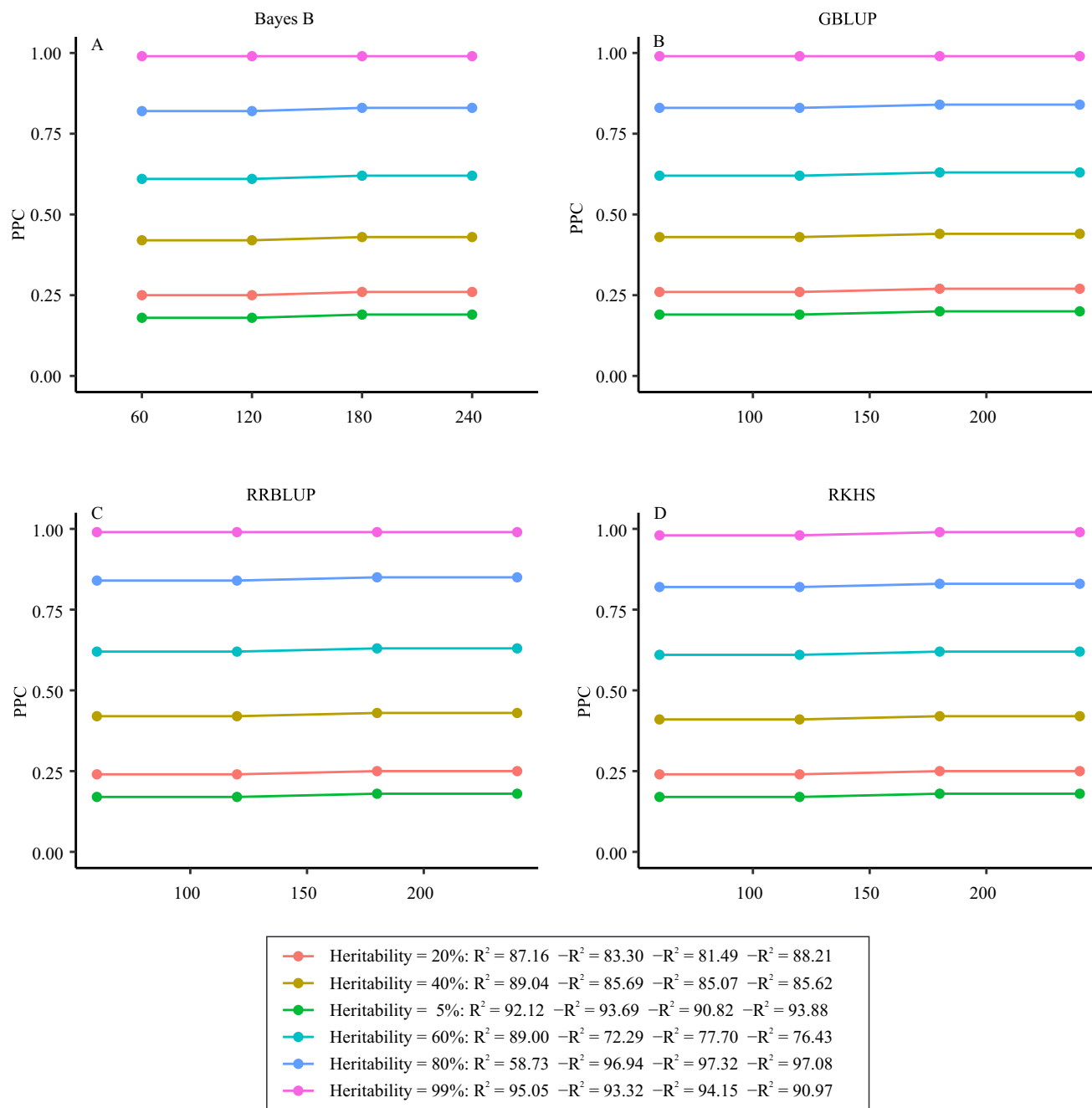


Figure 3. Phenotypic predictive ability (PPC) of the following four genomic selection methods in function of the number of QTL controlling the trait (60, 120, 180, and 240) and their respective coefficient of determination (R^2) values, evaluated in different heritability values: Bayesian estimation method B (A), genomic best linear unbiased prediction (B), reproducing kernel Hilbert spaces regression (C), and ridge-regression best linear unbiased prediction (D).

regardless of the genomic selection method used to estimate the genetic and phenotypic correlations.

According to Dekkers (2007), accuracy, also known as the genetic correlation between the true genetic value and the GEBV, is estimated by the correlation

between the phenotypic value and the GEBV divided by heritability root squared. The linear, quadratic, and cubic regression models were used to predict the genetic correlation in function of the phenotypic correlation. However, the R^2 evaluation of the regression models

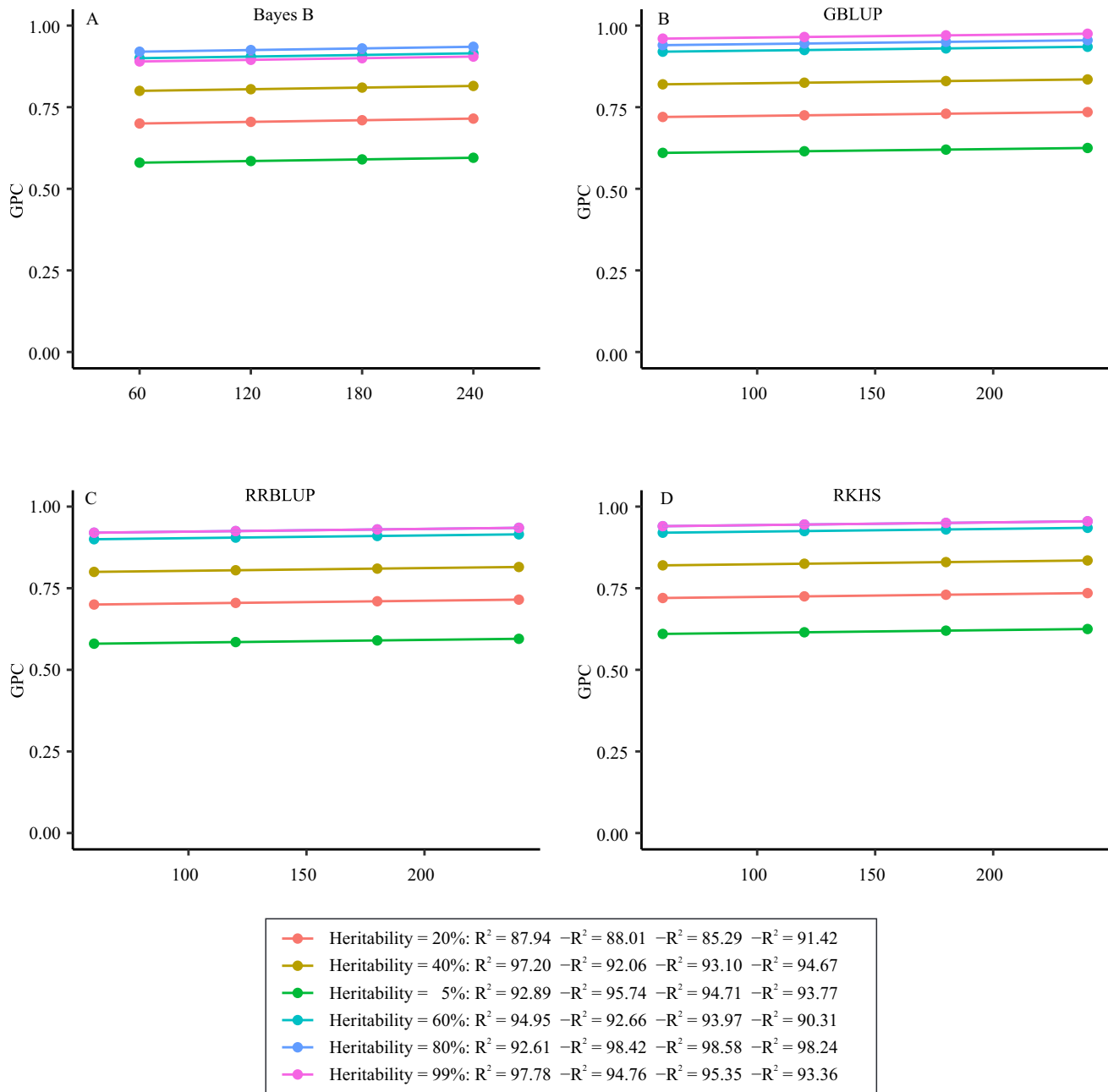


Figure 4. Genotypic predictive ability (GPC) of the following four genomic selection methods in function of the number of QTL controlling the trait (60, 120, 180, and 240) and their respective coefficient of determination (R^2) values, evaluated in different heritability values: Bayesian estimation method B (A), genomic best linear unbiased prediction (B), reproducing kernel Hilbert spaces regression (C), and ridge-regression best linear unbiased prediction (D).

Table 3. Coefficient of determination for prediction accuracy by phenotypic accuracy for traits controlled by 60 quantitative trait loci obtained for the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods.

Regression model	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
	Heritability = 5%				Heritability = 20%			
Linear	32.20	30.53	30.68	35.27	48.31	10.57	32.67	38.97
Quadratic	33.16	30.69	31.17	35.56	48.67	11.20	32.78	38.97
Cubic	34.10	36.09	31.20	35.70	49.80	13.97	33.68	40.29
	Heritability = 40%				Heritability = 60%			
Linear	29.00	35.41	61.54	54.33	51.08	61.65	40.73	34.34
Quadratic	29.03	37.49	61.69	55.73	51.15	61.67	41.17	34.37
Cubic	29.68	37.62	63.35	57.10	51.34	62.22	41.17	34.68
	Heritability = 80%				Heritability = 99%			
Linear	50.63	69.96	54.71	60.70	94.30	96.00	44.52	97.30
Quadratic	51.11	69.98	54.72	61.74	94.44	96.14	44.52	97.31
Cubic	51.11	70.08	54.72	61.79	94.44	96.14	44.52	97.31

Table 4. Coefficient of determination for prediction accuracy by phenotypic accuracy for traits controlled by 120 quantitative trait loci obtained for the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods.

Regression model	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
	Heritability = 5%				Heritability = 20%			
Linear	10.46	41.52	18.32	24.00	25.92	44.02	27.74	46.04
Quadratic	13.15	41.56	19.30	30.45	27.28	44.55	30.25	46.04
Cubic	13.23	47.76	19.71	35.70	27.46	44.75	30.25	46.85
	Heritability = 40%				Heritability = 60%			
Linear	21.10	17.80	52.50	11.62	36.73	43.84	34.77	52.93
Quadratic	23.67	25.27	57.68	12.09	37.29	49.81	35.52	52.97
Cubic	28.49	25.72	58.29	12.23	37.79	50.49	45.11	53.08
	Heritability = 80%				Heritability = 99%			
Linear	68.36	72.07	58.31	54.47	94.63	96.88	10.23	95.43
Quadratic	68.47	72.56	58.77	54.56	94.64	96.91	10.23	95.89
Cubic	69.01	72.63	59.44	59.04	94.64	96.91	10.23	95.89

Table 5. Coefficient of determination for prediction accuracy by phenotypic accuracy for traits controlled by 180 quantitative trait loci obtained for the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods.

Regression model	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
Heritability = 5%					Heritability = 20%			
Linear	51.23	21.41	24.94	27.04	32.00	44.90	35.76	30.18
Quadratic	51.31	21.58	24.97	27.13	33.06	44.94	36.49	38.49
Cubic	51.61	21.94	26.09	30.31	35.96	45.09	36.59	41.45
Heritability = 40%					Heritability = 60%			
Linear	39.73	35.04	36.26	19.38	45.74	49.68	44.41	51.75
Quadratic	39.84	35.41	36.26	19.56	46.41	53.45	56.03	51.75
Cubic	40.03	36.84	36.27	20.15	47.30	53.59	56.09	52.25
Heritability = 80%					Heritability = 99%			
Linear	45.01	52.35	26.63	63.75	95.96	95.00	40.83	95.24
Quadratic	45.01	52.52	26.69	65.35	95.97	95.00	40.83	95.29
Cubic	45.03	52.70	26.69	65.48	95.97	95.00	40.83	95.29

Table 6. Coefficient of determination for prediction accuracy by phenotypic accuracy for traits controlled by 240 quantitative trait loci obtained for the ridge-regression best linear unbiased prediction (RR-BLUP), Bayesian estimation method B (BB), reproducing kernel Hilbert spaces regression (RKHS), and genomic BLUP (GB) methods.

Regression model	RR-BLUP	BB	RKHS	GB	RR-BLUP	BB	RKHS	GB
Heritability = 5%				Heritability = 20%				
Linear	10.43	38.68	33.45	42.81	25.02	42.59	35.19	25.70
Quadratic	14.06	38.84	34.80	43.03	26.67	42.60	35.19	26.26
Cubic	14.43	40.94	37.03	43.94	28.06	43.27	38.07	26.65
Heritability = 40%				Heritability = 60%				
Linear	26.49	39.15	42.28	26.06	52.81	55.62	42.87	56.80
Quadratic	30.60	39.60	43.64	26.07	53.38	58.53	42.93	56.80
Cubic	33.41	42.57	43.65	26.12	53.54	63.81	43.37	58.06
Heritability = 80%				Heritability = 99%				
Linear	61.50	49.10	35.61	67.03	89.23	92.79	58.35	93.37
Quadratic	61.92	49.14	35.78	69.03	89.24	92.83	58.35	93.83
Cubic	61.92	49.14	35.78	69.09	89.24	92.83	59.08	93.83

revealed that the relationship between the genetic and phenotypic correlations cannot be explained by simple regression models (Tables 3, 4, 5, and 6).

For heritability values closer to 1, i.e., for a small environmental effect, the regression models explained more accurately the genetic correlation from the phenotypic correlation (Tables 3, 4, 5, and 6). This fact is explained by the relationship between heritability and correlation, in which the correlation of the phenotypic value with the genetic value is the square root of heritability.

The results obtained in the present study are an indicative that there is no linear relationship between genetic and phenotypic correlations when the heritability of the trait is lower than 80%. This means that nonlinear models, such as artificial neural networks, must be used to estimate more accurately the genetic correlation in function of the phenotypic correlation.

Conclusions

1. Heritability presents a positive linear relationship with the phenotypic predictive ability and a positive quadratic relationship with the genotypic predictive ability of the evaluated genomic selection methods.
2. The number of QTL controlling the trait has no relationship with the phenotypic and genotypic predictive abilities of the tested methods.

Acknowledgments

To Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), to Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), and to Fundação Arthur Bernardes (FUNARBE), for financial support; and to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), for financing, in part, this study (Finance Code 001).

References

- BHERING, L.L.; JUNQUEIRA, V.S.; PEIXOTO, L.A.; CRUZ, C.D.; LAVIOLA, B.G. Comparison of methods used to identify superior individuals in genomic selection in plant breeding. **Genetic Molecular Research**, v.14, p.10888-10896, 2015. DOI: <https://doi.org/10.4238/2015.september.9.26>.
- CRUZ, C.D. Genes: a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum. Agronomy**, v.35, p.271-276, 2013. DOI: <https://doi.org/10.4025/actasciagron.v35i3.21251>.
- DE LOS CAMPOS, G.; GIANOLA, D.; ROSA, G.J.M.; WEIGEL, K.A.; CROSSA, J. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. **Genetics Research**, v.92, p.295-308, 2010. DOI: <https://doi.org/10.1017/S0016672310000285>.
- DE MORI, G.; CIPRIANI, G. Marker-assisted selection in breeding for fruit trait improvement: a review. **International Journal of Molecular Sciences**, v.24, art.8984, 2023. DOI: <https://doi.org/10.3390/ijms24108984>.
- DEKKERS, J.C.M. Prediction of response to marker-assisted and genomic selection using selection index theory. **Journal of**

- Animal Breeding and Genetics**, v.124, p.331-341, 2007. DOI: <https://doi.org/10.1111/j.1439-0388.2007.00701.x>.
- DESTA, Z.A.; ORTIZ, R. Genomic selection: genome-wide prediction in plant improvement. **Trends Plant Science**, v.19, p.592-601, 2014. DOI: <https://doi.org/10.1016/j.tplants.2014.05.006>.
- HEFFNER, E.L.; JANNINK, J.-L.; IWATA, H.; SOUZA, E.; SORRELLS, M.E. Genomic selection accuracy for grain quality traits in biparental wheat populations. **Crop Science**, v.51, p.2597-2606, 2011. DOI: <https://doi.org/10.2135/cropsci2011.05.0253>.
- HESLOT, N.; YANG, H.-P.; SORRELLS, M.E.; JANNINK, J.-L. Genomic selection in plant breeding: a comparison of models. **Crop Science**, v.52, p.146-60, 2012. DOI: <https://doi.org/10.2135/cropsci2011.06.0297>.
- LAMICHHANE, S.; THAPA, S. Advances from conventional to modern plant breeding methodologies. **Plant Breeding and Biotechnology**, v.10, p.1-14, 2022. DOI: <https://doi.org/10.9787/PBB.2022.10.1.1>.
- MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v.157, p.1819-1829, 2001. DOI: <https://doi.org/10.1093/genetics/157.4.1819>.
- ORNELLA, L.; SINGH, S.; PEREZ, P.; BURGUEÑO, J.; SINGH, R.; TAPIA, E.; BHAVANI, S.; DREISIGACKER, S.; BRAUN, H.-J.; MATHEWS, K.; CROSSA, J. Genomic prediction of genetic values for resistance to wheat rusts. **The Plant Genome**, v.5, p.136-148, 2012. DOI: <https://doi.org/10.3835/plantgenome2012.07.0017>.
- R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2017.
- ROBERT, P.; AUZANNEAU, J.; GOUDEMAND, E.; OURY, F.-X.; ROLLAND, B.; HEUMEZ, E.; BOUCHET, S.; LE GOUIS, J.; RINCENT, R. Phenomic selection in wheat breeding: identification and optimisation of factors influencing prediction accuracy and comparison to genomic selection. **Theoretical and Applied Genetics**, v.135, p.895-914, 2022. DOI: <https://doi.org/10.1007/s00122-021-04005-8>.
- SONG, L.; WANG, R.; YANG, X.; ZHANG, A.; LIU, D. Molecular markers and their applications in marker-assisted selection (MAS) in bread wheat (*Triticum aestivum* L.). **Agriculture**, v.13, art.642, 2023. DOI: <https://doi.org/10.3390/agriculture13030642>.
- VOSS-FELS, K.P.; COOPER, M.; HAYES, B.J. Accelerating crop genetic gains with genomic selection. **Theoretical Applied Genetics**, v.132, p.669-686, 2019. DOI: <https://doi.org/10.1007/s00122-018-3270-8>.
- XU, Y.; CROUCH, J.H. Marker-assisted selection in plant breeding: from publications to practice. **Crop Science**, v.48, p.391-407, 2008. DOI: <https://doi.org/10.2135/cropsci2007.04.0191>.
- XU, Y.; MA, K.; ZHAO, Y.; WANG, X.; ZHOU, K.; YU, G.; LI, C.; LI, P.; YANG, Z.; XU, C.; XU, S. Genomic selection: a breakthrough technology in rice breeding. **The Crop Journal**, v.9, p.669-677, 2021. DOI: <https://doi.org/10.1016/j.cj.2021.03.008>.
- ZARGAR, S.M.; RAATZ, B.; SONAH, H.; NAZIR, M.; BHAT, J.A.; DAR, Z.A.; AGRAWAL, G.K.; RAKWAL, R. Recent advances in molecular marker techniques: insight into QTL mapping, GWAS and genomic selection in plants. **Journal of Crop Science and Biotechnology**, v.18, p.293-308, 2015. DOI: <https://doi.org/10.1007/s12892-015-0037-5>.
- ZHAO, Y.; GOWDA, M.; WÜRSCHUM, T.; LONGIN, C.F.H.; KORZUN, V.; KOLLERS, S.; SCHACHSCHNEIDER, R.; ZENG, J.; FERNANDO, R.; DUBCOVSKY, J.; REIF, J.C. Dissecting the genetic architecture of frost tolerance in Central European winter wheat. **Journal of Experimental Botany**, v.64, p.4453-4460, 2013. DOI: <https://doi.org/10.1093/jxb/ert259>.
- ZHONG, S.; DEKKERS, J.C.M.; FERNANDO, R.L.; JANNINK, J.-L. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. **Genetics**, v.182, p.355-364, 2009. DOI: <https://doi.org/10.1534/genetics.108.098277>.