

Effects of Marker Density, Number of Quantitative Trait Loci and Heritability of Trait on Genomic Selection Accuracy

Research Article

F. Alanoshahr^{1*}, S.A. Rafat¹, R. Imany-Nabiyyi², S. Alijani¹ and C. Robert Granie³

¹ Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

² Department of Animal Science, Faculty of Mathematical Science, University of Tabriz, Tabriz, Iran

³ INRA-INPT-ENSAT-INPT-ENVT, Université de Toulouse, UMR 1388 GenPhySE, Castanet Tolosan, France

Received on: 30 May 2016

Revised on: 11 Sep 2016

Accepted on: 15 Sep 2016

Online Published on: Dec 2017

*Correspondence E-mail: f.Alanoshahr@gmail.com

© 2010 Copyright by Islamic Azad University, Rasht Branch, Rasht, Iran

Online version is available on: www.ijas.ir

ABSTRACT

The success of genomic selection mainly depends on the extent of linkage disequilibrium (LD) between markers and quantitative trait loci (QTL), number of QTL and heritability (h^2) of the traits. The extent of LD depends on the genetic structure of the population and marker density. This study was conducted to determine the effects of marker density, level of heritability, number of QTL, and to compare the accuracy of predicting breeding values using two diverse approaches: GBLUP and BayesA using simulated data under two different distributions of the QTL effect. Thereby, three traits (milk production, carcass weight and mature body weight) were simulated with the heritability of 0.10, 0.30 and 0.50, respectively; for each ovine animal, a genome with three chromosomes, 100 cM each. Three different marker densities was considered (1000, 2000 and 3000 markers) and the number of QTL was assumed to be either 100, 200 or 300. Data were simulated with two different distributions of the QTL effect which were uniform and gamma ($\alpha=1.66$ and $\beta=0.4$) the marker density, number of the QTL, the QTL effect distributions and heritability levels significantly affected the accuracy of genomic breeding values ($P<0.05$). The BayesA produced estimates with greater accuracies in traits influenced by a low number of the QTL and with the gamma QTL effects distribution. Based on the findings of this simulation, heritability, as well as dense marker panels, aiming to increase the level of LD between the markers and QTL, is likely to be needed for successful implementation of the genomic selection.

KEY WORDS BayesA, GBLUP, genomic breeding value, heritability, marker density, quantitative trait loci.

INTRODUCTION

Genomic selection is a method of marker-assisted selection based on linkage disequilibrium (LD) that potentially explores all QTL in the genome (Meuwissen *et al.* 2001). Selection based on the genome wide distributed markers estimated breeding values (MEBVs) resulted in increased genetic progress, due to improvement in the accuracy of estimations of MEBVs, reduction in the generation interval (Meuwissen *et al.* 2001) and reduction in inbreeding rates,

due to emphasis on MEBVs rather than family information (Daetwyler *et al.* 2007). The accuracy in obtaining MEBVs determines the success rate in the breeding programs. The reason of this process is that whenever marker density is high enough, most QTL will be in high LD with some markers and estimates of marker effects will lead to accurate predictions of genetic merit for a trait. Despite this, the amount of information to be analyzed in this situation poses new challenges from statistical and computational points of view.

The accuracy of MEBV depends on genetic architecture of the trait, number of the QTL, marker density panels, the heritability of the trait, the size of the training population, the distribution of QTL variance, the method used to estimate marker effects and the LD between markers and QTL (Meuwissen *et al.* 2001; Shirali *et al.* 2012). There are two main approaches in the genomic selection for estimating breeding values. The first approach assumes that all single nucleotide polymorphisms (SNPs) have effects on the trait variance and the second approach assumes that only some SNPs contribute to the trait variance. In the first approach, genomic best linear unbiased prediction (GBLUP) methods including a form of ridge regression (Meuwissen *et al.* 2001) are applied, which instead of a pedigree relationship matrix, the marker relationship matrix is used (NejatiJavaremi *et al.* 1997). The second approach assumes that a limited number of SNPs contribute to the trait variances and that among these affecting SNPs, only few of them make large contributions to trait variance and the other have small contributions. In this approach, Bayesian methods (e.g., BayesA, BayesB, BayesC and Bayesian Lasso) have usually been used (Tibshirani, 1996; Meuwissen *et al.* 2001).

As it has been mentioned, one of the main statistical challenges in the genomic selection is that in the most cases the number of parameters to be estimated, i.e., marker effects, is more than the number of records. Many innovative statistical methods have been proposed to overcome this challenge, including ridge regression (Whittaker *et al.* 2000), Bayesian methods such as BayesA and BayesB (Meuwissen *et al.* 2001), BayesC and Bayes C π (Habier *et al.* 2011), Bayesian Least Absolute Shrinkage and Selection Operator (LASSO) (De los Campos *et al.* 2009), and nonparametric kernel methods (Gianola and van Kaam, 2008). The Bayesian methods use a prior for the QTL effect distribution and another prior for the number of the QTL (Meuwissen *et al.* 2001). However, the true distribution of the QTL effects is unknown for many quantitative traits. Goddard (2008) found greater accuracies by using a gamma (1.66, 0.4) prior distribution for the QTL effects compared to a normal prior distribution. Briefly, the GBLUP assumption is that the genetic model is an infinitesimal model and all SNPs have effects on the trait of interest, while the BayesA approach assumes all SNP to have some effects, however, assumed that some of the SNPs are in LD with QTL of moderate to large effect. However, BayesB approach assumes many SNPs are in the genomic regions where there are no QTL and thus have zero effects, whilst small proportions of SNPs are in LD with the QTL and consequently do have an effect.

Factors affecting the accuracy of the prediction of the genotypes are largely unknown. At present it is unknown

how dense markers and QTL number need to be, particularly if they vary in information content. Therefore, the main objective of this study was to investigate the effects of marker density, number of the QTL, heritability levels and the QTL effect distribution on the accuracy of MEBVs using the GBLUP and BayesA approaches with simulated populations.

MATERIALS AND METHODS

Simulation

The populations were simulated using the QMSim® software (Sargolzaei and Schenkel, 2009) based on forward-in-time process. A genome consisted of three chromosomes with a length of 100 cM was simulated 1000, 2000 and 3000 SNPs were equally spaced over the chromosomes. Three different numbers of QTL (100, 200 and 300) were considered and QTLs were uniformly distributed over the chromosomes. One hundred individuals, including 50 males and 50 females, were simulated for the base population (zero generation). These individuals were assumed to be biallelic for both SNPs and QTL with allele frequencies equal to 0.50 (Table 1). For the first generation, one male and one female were randomly chosen from the base population as parents. The parent's gametes were simulated assuming LD based on the Haldane mapping function (Haldane, 1919) to generate the recombinant gametes and were randomly combined to create the individual. The first generation structure was followed through to the 50th generation of random mating to make linkage disequilibrium populations. The occurred recombination in the chromosome had a poisson distribution. For each generation, LD was measured using r^2 which was the average LD of all SNPs. Subsequent to the LD populations ten more generations (51 to 60) were constructed.

The base population consisted of 1000 unrelated animals (500 males and 500 females). In this study, generations 51 and 52 were assumed as training population and the other generations (53 to 60) as validation populations. In simulating training and validation populations, three QTL numbers (100, 200 and 300), three marker densities (1000, 2000 and 3000) and three heritability levels of 0.10, 0.30 or 0.50 were assumed to be influencing the trait of interest. Furthermore, two different assumed distributions for the QTL effect were uniform and gamma (1.66, 0.4). Overall these assumptions for simulations generated traits in this study that had different genetic architectures. The mutation rate of the markers and QTLs was assumed 2.5×10^{-5} per locus per generation (Solberg *et al.* 2008).

The true breeding value (TBV) of each individual was equal to the sum of the QTL allele substitution effects, assuming only additive QTL effects.

Table 1 Population structure and simulation parameters

Parameter	Value
Number of chromosomes	3
Genome length	300cM
Effective population size (Ne)	50, 100 and 200
Numbers of QTL	100, 200 and 300
Number of SNP markers	1000, 2000 and 3000
QTL effects distributions	Uniform and gamma (1.66; 0.4)
Heritability	0.1, 0.3 and 0.5
Training set	All individuals of generation 51 and 52
Validation set	All individuals of generations 53 to 60

Phenotypes were generated by adding residuals, randomly drawn from a normal distribution with mean equal to zero, to the TBVs. For all scenarios, 10 replicates were simulated.

Linkage disequilibrium

The LD measure r^2 (square of the correlation of alleles at two loci) was used for measuring LD (Hill and Robertson, 1968):

$$r^2 = \frac{D^2}{f(A) * f(a) * f(B) * f(b)}$$

Where:

$$D = f(AB) - f(A).f(B).$$

$f(AB)$, $f(A)$, $f(a)$, $f(B)$, $f(b)$: observed frequencies of haplotypes AB and of alleles A, a, B, b, respectively.

Estimating the breeding values

Two methods including GBLUP and BayesA, were used to estimate the QTL and SNPs effects, and the genomic breeding values. The main difference between these two approaches is in their assumptions regarding genetic models of the trait. The genomic estimated breeding value for individuals in validation generations for two GBLUP and BayesA methods were predicted using the model:

$$GEBV = \sum_i^n X_{i,t} \hat{\beta}_i$$

Where:

n : number of chromosomes across the genome.

X_i : design matrix which refers to the individual genotypes for the chromosome i .

g_i : vector of SNP effects in chromosome i .

GBLUP method

The GBLUP approach was based on the simple mixed model and assumed that all SNPs had equal effects on genetic variance of the considered trait. In the GBLUP approach, the following model was applied using ASReml (Gilmour *et al.* 1995):

$$y = \mu 1_n + X_i g_i + e$$

Where:

y : vector of the phenotypic values.

μ : overall mean.

1_n : vector of n ones.

n : number of records.

X_i is the design matrix for the i^{th} SNP.

g_i : represents the additive genetic effects of the i^{th} SNP.

e : vector of residual error with normal distribution.

The additive genetic effects of SNPs (g) were assumed to have a normal distribution $N(0, \sigma_g)$ where, g was the realized relationship matrix for all loci. The g was calculated based on the identical-by-state probabilities between a pair of individuals for all individuals in the training and validation populations. The total allelic relationship between each pair of individuals was calculated based on the method of Nejati-Javaremi *et al.* (1997). The mixed model equation to obtain breeding values is:

$$\begin{bmatrix} 1_n & 1_n & 1_n & X \\ X & 1_n & X & X + I\alpha \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} 1_n y \\ X y \end{bmatrix}$$

Where:

$$\alpha = \frac{\sigma_e^2}{\sigma_g^2}$$

σ_g^2 : variance common to each marker effect.

σ_e^2 : residual variance and I is the identity matrix.

BayesA methods

In this model like GBLUP, all SNPs are assumed to have small effects, however, assumed that some of the SNPs are in LD with the QTL of moderate to large effects. The SNP effects sampled from a normal distribution with the variance for each SNP sampled from an inverse scaled chi-square distribution (Meuwissen *et al.* 2001). The BayesA method was performed using the model:

$$f(e|\sigma_e^2) \sim N(0, I\sigma_e^2), f(\sigma_e^2|a, b) \sim U(a, b)$$

$$f(\sigma_{g_j}^2|v, s^2) \sim \chi^{-2}(v, s^2)$$

$$v = 4.2, \hat{\sigma}_{SNP}^2 = \frac{\hat{\sigma}_a^2}{(1-\pi) \sum_{m=1}^M 2p_m q_m}$$

$$s^2 = \frac{\hat{\sigma}_{SNP}^2(v-2)}{v}$$

Where:

S: scale parameter.

v: number of degrees of freedom.

For each analysis, a Markov chain Monte Carlo (MCMC) with 210000 cycles with WinBUGS 1.4 software ran and the first 10000 cycles were discarded as burn-in period. Estimates at every 5th iteration were sorted as a sample, resulting in a total 40000 samples (Meuwissen *et al.* 2001).

Comparison of the methods to estimate breeding values

The effects of heritability levels, marker density panels, and the number of QTLs on the accuracy of genomic predictions were evaluated using PROC GLM, and the average accuracies of GEBV were compared using the least squares means (LSM) procedure at $P < 0.05$ (SAS, 2003). The correlation between the GEBV and true genomic breeding value (TGBV) was used as measure of accuracy.

RESULTS AND DISCUSSION

In current simulation analysis, calculated average LD values between all SNPs (r^2) in the last generation of the LD population (generation 50) were 0.191 ± 0.011 . This indicated that 87% of the expected LD had been achieved in this simulation. The expected LD based on Sved (1971) formula was 0.210.

In this study, the genomic accuracy, the correlations between the TBVs and GEBVs, for different marker densities (1000, 2000 and 3000), different number of the QTLs (100, 200 and 300), different levels of heritability (0.10, 0.30, and 0.50) with the two QTL effect distributions uniform and gamma are shown in Tables 2 and 3, respectively.

Accuracy under different numbers of QTL

The results showed that increasing the number of QTLs from 100 to 300 could decrease the average genomic accuracy from 0.708 to 0.687 in GBLUP and 0.713 to 0.689 in BayesA with the uniform QTL effect distribution (Table 2).

Also in the gamma QTL effect distribution, the average genomic accuracy from 0.719 to 0.697 in GBLUP and

0.725 to 0.704 in BayesA decreased (Table 3). The accuracy of BayesA was the greatest at low number of the QTL (100) and then decreased with increasing number of the QTL. However, GBLUP has an advantage in comparison to BayesA at high number of the QTL (300), but this advantage decreased by number of QTL decreased (Figure 1).

Accuracy under different marker density

In this study for trait with low heritability (0.1), increasing the marker density from 1000 to 2000 increased the average genomic accuracy with both two methods and two QTL effect distributions.

But by increasing the marker density to 3000, not only no increase the genomic accuracy but also slightly decreased the genomic accuracy was observed. While in traits with moderate and high heritability (0.3 and 0.5) increasing the marker density from 1000 to 3000 increased the average genomic accuracy with both two methods and two QTL effect distributions. Among the two studied models, the greatest amount of genomic accuracy was 0.797 (0.01) that obtained by BayesA in trait with heritability by 0.5, low QTL number (100), high marker density and the gamma QTL effect distribution.

Accuracy under different heritability levels

Under all situations, the average genomic accuracy increased when heritability values increased from 0.10 to 0.50, ($P < 0.05$). Increasing the level of heritability from 0.10 to 0.50 increased the average genomic accuracy from 0.674 (0.012) to 0.733 (0.01) in GBLUP and 0.677 (0.02) to 0.74 (0.01) in BayesA with uniform QTL effect distribution. Also such changes observed in genomic accuracy with gamma QTL effect distribution as accuracy increased from 0.684 (0.01) to 0.738 (0.01) in GBLUP and 0.691 (0.01) to 0.749 (0.01) in BayesA.

In general, in the same number of QTL and density of markers in both uniform and gamma QTL effect distribution, estimated accuracies of genomic breeding values for traits with high heritability, were greater than the moderate and low heritability traits respectively.

The results showed that, when the heritability of a trait was greater, the estimated accuracies of genomic breeding values were greater.

Accuracy under different QTL effect distribution

Traits with uniform QTL effect distribution (Table 2) and gamma QTL effect distribution (Table 3), verify that accuracy of genomic breeding values in the gamma distribution provides better gene effects to uniform distribution. Also, when the distribution of the gene effect was gamma, the genomic accuracy of Bayesian method was greater than GBLUP.

Table 2 The estimated genomic accuracy (SE) for three marker densities, levels of quantitative trait loci (QTL) and heritability by uniform QTL variance

		Statistical methods					
Heritability	N _{QTL}	GBLUP			BayesA		
		Marker density					
		1000	2000	3000	1000	2000	3000
0.1	100	0.674 (0.01)	0.685 (0.01)	0.679 (0.01)	0.680 (0.02)	0.691 (0.02)	0.685 (0.02)
	200	0.662 (0.01)	0.681 (0.01)	0.675 (0.01)	0.673 (0.02)	0.684 (0.02)	0.680 (0.02)
	300	0.654 (0.01)	0.672 (0.01)	0.671 (0.01)	0.652 (0.02)	0.659 (0.02)	0.669 (0.02)
0.3	100	0.685 (0.01)	0.688 (0.01)	0.694 (0.01)	0.687 (0.02)	0.693 (0.02)	0.705 (0.02)
	200	0.676 (0.01)	0.679 (0.01)	0.687 (0.01)	0.677 (0.02)	0.686 (0.02)	0.765 (0.02)
	300	0.669 (0.01)	0.675 (0.01)	0.682 (0.01)	0.660 (0.02)	0.673 (0.02)	0.679 (0.02)
0.5	100	0.729 (0.01)	0.756 (0.01)	0.783 (0.01)	0.732 (0.01)	0.758 (0.01)	0.789 (0.01)
	200	0.718 (0.01)	0.729 (0.01)	0.735 (0.01)	0.725 (0.01)	0.733 (0.01)	0.745 (0.01)
	300	0.709 (0.01)	0.715 (0.01)	0.722 (0.01)	0.706 (0.01)	0.711 (0.01)	0.718 (0.01)

GBLUP: genomic best linear unbiased prediction.

Table 3 The estimated genomic accuracy (SE) for three marker densities, levels of quantitative trait loci (QTL) and heritability by gamma the QTL variance

		Statistical methods					
Heritability	N _{QTL}	GBLUP			BayesA		
		Marker density					
		1000	2000	3000	1000	2000	3000
0.1	100	0.682 (0.01)	0.698 (0.01)	0.695 (0.01)	0.691 (0.01)	0.704 (0.01)	0.698 (0.01)
	200	0.673 (0.01)	0.684 (0.01)	0.679 (0.01)	0.682 (0.01)	0.697 (0.01)	0.686 (0.01)
	300	0.667 (0.01)	0.683 (0.01)	0.671 (0.01)	0.661 (0.01)	0.680 (0.01)	0.667 (0.01)
0.3	100	0.694 (0.01)	0.703 (0.01)	0.717 (0.01)	0.698 (0.01)	0.709 (0.01)	0.720 (0.01)
	200	0.682 (0.01)	0.693 (0.01)	0.707 (0.01)	0.689 (0.01)	0.700 (0.01)	0.711 (0.01)
	300	0.673 (0.01)	0.688 (0.01)	0.697 (0.01)	0.671 (0.01)	0.683 (0.01)	0.695 (0.01)
0.5	100	0.733 (0.01)	0.757 (0.01)	0.791 (0.01)	0.744 (0.01)	0.763 (0.01)	0.797 (0.01)
	200	0.720 (0.01)	0.733 (0.01)	0.747 (0.01)	0.731 (0.01)	0.746 (0.01)	0.759 (0.01)
	300	0.714 (0.01)	0.722 (0.01)	0.730 (0.01)	0.712 (0.01)	0.719 (0.01)	0.725 (0.01)

GBLUP: genomic best linear unbiased prediction.

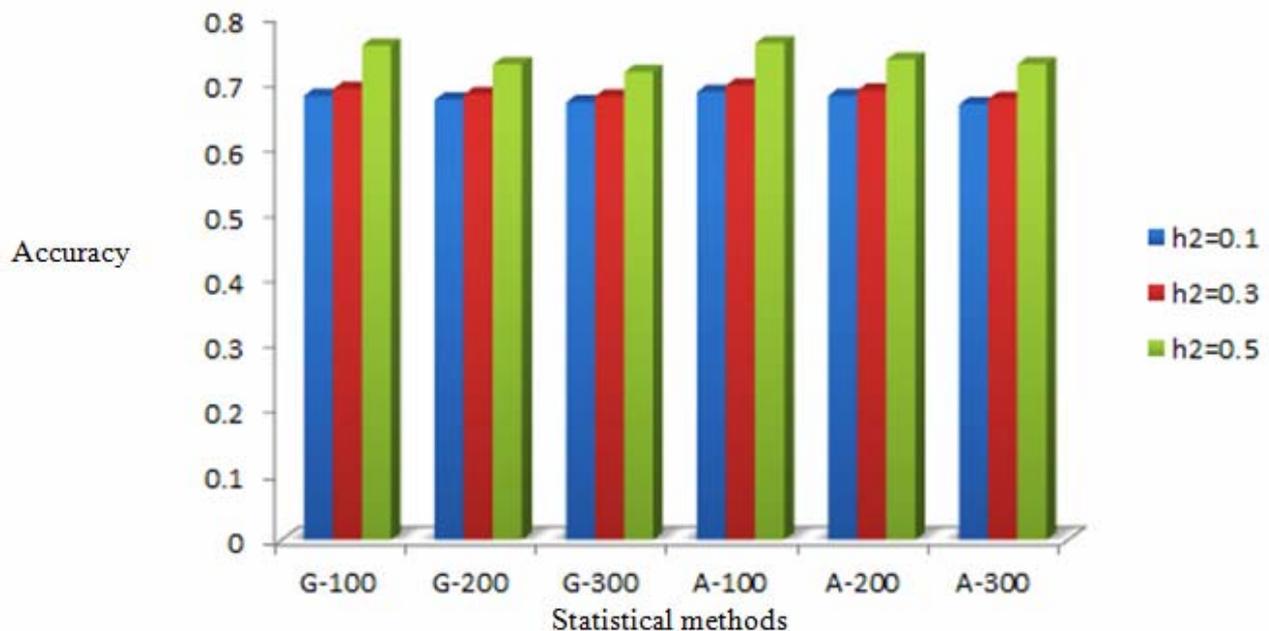


Figure 1 Genomic accuracy of two methods including GBLUP (G-100, G-200 and G-300 = GBLUP with 100, 200 and 300 QTL) and BayesA (A-100, A-200 and A-300 = BayesA with 100, 200 and 300 QTL) with three number of the QTL (100, 200 and 300) in same marker density and the QTL effect distribution

The highest accuracy of genomic breeding value obtained 0.789 by BayesA method with uniform distribution of QTL effect in trait with high heritability (0.5), and this accuracy on the same level with the gamma distribution of QTL effect was 0.797, and this difference was statistically significant ($P < 0.05$). The lowest accuracy of genomic breeding value in both variance distributions achieved by GBLUP method. By changing the uniform distribution of variance to the gamma in GBLUP, it was seen a slight increase in accuracy genomic breeding value, but this increase was not statistically significant ($P > 0.05$).

The maximum accuracy of BayesA estimates was achieved for the lowest number of the QTL. This indicates that BayesA has an advantage over GBLUP for analysis of traits that are influenced by a low number of the QTL. The results of the current study are in agreement with [Daetwyler et al. \(2010\)](#) who found a decrease in the accuracy with an increase in the number of the QTL. By increasing the number of the QTL for a trait, the average variance of each the QTL for the trait of interest will decrease and the estimation of the QTL effect will be less accurate. With uniform QTL effect distribution, by increasing the number of the QTL the proportional contribution of each the QTL on the trait will be very low and therefore some of their effects will be missed and missing heritability will be increased.

However, in this study GBLUP has an advantage over BayesA at high number of the QTL (300), but this advantage decreased by reducing the number of the QTL. Studies in this context have shown that the number of QTL affecting the trait is high, GBLUP way similar or even better than Bayesian methods ([Daetwyler et al. 2010](#)).

In addition, in the current study the gamma distributions of the QTL effects resulted in better accuracy in both methods. [Shirali et al. \(2012\)](#) also reported better accuracy using BayesC estimation for gamma distribution in the QTL variance. When the distribution of the gene effects is gamma, some genes have a major effects and a high percentage of genes are close to zero impact. So the Bayesian methods are better than non-Bayesian methods. These effects can be due to two possible reasons. First, the prior, the QTL effect and the QTL variance had all gamma distributions. As a result, BayesA would have a better estimation of any SNP effect. Second, the gamma distribution captures the QTL with very high effects compared to a normal and uniform distribution, resulting in more accurate estimation of GEBVs for traits which are influenced by a number of the QTL with high effects.

In this study in trait with low heritability (0.1), increasing the marker density from 1000 to 2000 increased the average genomic accuracy with two methods and the two QTL effect distributions. But by increasing marker density to 3000,

not only did not increase the genomic accuracy but also slightly decreased the genomic accuracy. While in traits with moderate and high heritability (0.3 and 0.5) increasing the marker density from 1000 to 3000 increased the average genomic accuracy with two methods and two QTL effect distributions. Increase the accuracy of genomic breeding values by increasing marker density to 2000 can be resulted to increase the linkage disequilibrium between markers and QTL.

But reducing the accuracy with increase the marker density up to 3000 markers on each chromosome can be due to the increased number of markers or increase the number of unknown variables (the marker effects) and deficiency accurate estimation of the marker effects for traits with low heritability.

In traits with lower heritability, the correlation between the phenotype and genetic value will be lower and estimates of the marker effects can be done with less precision ([Habier et al. 2011](#)). The results of this study were agreement with the results of [Solberg et al. \(2008\)](#) and [Meuwissen et al. \(2001\)](#).

High heritability, meaning that the more additive variance to phenotypic variance ratio, leads to more of the genes role in the development diversity of traits and therefore, more accurate estimation of the marker effects. The accuracy genomic breeding value with high heritability traits is greater than the low heritability traits. The reason is clear because when the heritability of trait is high, individual phenotypic value closer genetic value and thus genomic breeding values are estimated to be more accurate ([Goddard, 2008](#)). In current study, in comparison two methods GBLUP and BayesA, BayesA method was more accurate. Many studies have shown that Bayesian methods are more accurate in comparison with BLUP which is consistent with the present results ([Calus et al. 2008](#); [Solberg et al. 2008](#)). One weakness of the GBLUP method is to consider the same proportion of the variance for all markers in the genomic prediction of breeding values.

While in the Bayesian method based on the prior distribution, be assigned different weights to the marker density. In GBLUP method, equal variance in all markers is considered and it is no longer necessary to have preliminary information on the variance of the marker effects (what is needed in the Bayesian approach). This method is simpler than the Bayesian method and requires less computation.

[Goddard \(2008\)](#) derived accuracies for GBLUP and Bayesian methods. The result of this study showed that greater accuracy can be obtained with Bayesian methods because this method better takes into account the variable contribution of the individual QTL. Bayesian methods should be preferred over GBLUP.

CONCLUSION

The extent of LD have major impact on the accuracy of MEBV. Based on the findings of this simulation study, low QTL number, as well as high dense marker panels, aiming to increase the level of LD between markers and the QTL, will likely be needed for successful implementation of the genomic selection. Gains in accuracy by using MEBV showed more advantage for traits with a moderate to high heritability. By using a dense marker map covering all chromosomes, it is possible to accurately estimate the breeding value of animals that have no phenotypic record of their own. The GBLUP method of analysis was as good as the BayesA method for traits influenced by a high number of QTL. BayesA produced estimates with higher accuracies in traits influenced by low number of QTL and with a gamma QTL effect distribution. the greater accuracy can be obtained with Bayesian methods because this method better takes into account the variable contribution of the individual QTL. Based on this, Bayesian Methods should be preferred over GBLUP.

ACKNOWLEDGEMENT

The authors thank C. Robert Granie for encouraging us to write this article and for comments provided on earlier versions of this manuscript.

REFERENCES

- Calus M.P.L., Meuwissen T.H.E., De Roos A.P.W. and Veerkamp R.F. (2008). Accuracy of genomic selection using different methods to define haplotypes. *Genetics*. **178**(1), 553-561.
- 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.
- Daetwyler H.D., Villanueva B., Bijma P. and Woolliams J.A. (2007). Inbreeding in genome-wide selection. *J. Anim. Breed. Genet.* **124**, 369-376.
- De los Campos G., Naya H., Gianola D., Crossa J., Legarra A., Manfredi E., Weigel K. and Cotes J.M. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*. **182**, 375-385.
- Gianola D. and van Kaam J. (2008). Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics*. **178**(4), 2289-2303.
- Gilmour A.R., Thompson R. and Cullis B.R. (1995). Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics*. **51**, 1440-1450.
- Goddard M. (2008). Genomic selection prediction of accuracy and maximisation of long term response. *Genetics*. **136**, 245-257.
- Habier D., Fernando R.L., Kizilkaya K. and Garrick D.J. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinform.* **12**, 186-193.
- Haldane J.B.S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. *Genetics*. **8**, 299-309.
- Hill W.G. and Robertson A. (1968). Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* **38**, 226-231.
- Meuwissen T.H.E., Hayes B.J. and Goddard M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. **157**, 321-322.
- Nejati Javaremi A., Smith C. and Gibson P.J. (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *J. Anim. Sci.* **75**, 1738-1745.
- Sargolzaei M. and Schenkel F.S. (2009). QMSim: a large-scale genome simulator for livestock. *Bioinformatics*. **25**, 680-681.
- SAS Institute. (2003). SAS[®]/STAT Software, Release 9.1. SAS Institute, Inc., Cary, NC. USA.
- Shirali M., Miraei-Ashtiani S.R., Pakdel A., Haley C. and Pong-Wong R. (2012). Comparison between Bayesc and GBLUP in estimating genomic breeding values under different QTL variance distributions. *Iranian J. Anim. Appl. Sci.* **43**, 261-268.
- Solberg T.R., Sonesson A.K., Woolliams J.A. and Meuwissen T.H.E. (2008). Genomic selection using different marker types and densities. *J. Anim. Sci.* **86**, 2447-2454.
- Sved J.A. (1971). Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor. Popul. Biol.* **2**, 125-141.
- Tibshirani R. (1996). Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. B Met.* **58**, 267-288.
- Whittaker J.C., Thompson R. and Denham M.C. (2000). Marker-assisted selection using ridge regression. *Genet. Res.* **75**(2), 249-252.