IJAS Iranian Journal of Applied Animal Science

Assessment of Alternative Single Nucleotide Polymorphism (SNP) Weighting Methods for Single-Step Genomic Prediction of Traits with Different Genetic Architecture

Research Article

S. Moghaddaszadeh-Ahrabi[1*] and M. Bazrafshan[2]

[1] Department of Animal Science, Faculty of Agriculture and Natural Resources, Islamic Azad University, Tabriz Branch, Tabriz, Iran
[2] Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Pakdasht, Tehran, Iran

Received on: 23 Jul 2022
Revised on: 2 Dec 2022
Accepted on: 8 Dec 2022
Online Published on: Jun 2023

*Correspondence E-mail: s.moghaddaszadeh@iaut.ac.ir
© 2010 Copyright by Islamic Azad University, Rasht Branch, Rasht, Iran
Online version is available on: www.ijas.ir

ABSTRACT

We investigated the prediction accuracy and bias of single-step genomic BLUP (ssGBLUP) with or without weights for single-nucleotide polymorphisms (SNPs). The SNP weights were calculated using population Fixation Index (WssGBLUP$_{FST}$) and a nonlinear method called nonlinearA (WssGBLUP$_{NLA}$). The results of these two weighted methods were compared with a non-weighted method. The individuals of the reference population were sorted based on their estimated breeding values and the top 5% and bottom 5% of individuals based on their estimated breeding values (EBVs) were considered as subpopulations 1 and 2. The FST values for all SNPs between subpopulations 1 and 2 were scaled between zero and one and used as weights. The prediction accuracy and bias of predictions in WssGBLUP$_{FST}$, WssGBLUP$_{NLA}$ and ssGBLUP methods were compared considering varying the numbers of quantitative trait locus (QTL) (10, 50 and 500), heritability (0.1 and 0.4) and size of reference population (1500, 5000 and 12500). In 10 and 50 QTL, both weighting methods outperformed regular ssGBLUP and with simulation, WssGBLUP$_{FST}$ outperformed WssGBLUP$_{NLA}$. By increasing the number of QTL to 500 QTL, the WssGBLUP$_{FST}$ was no longer superior to WssGBLUP$_{NLA}$ and ssGBLUP. Our results suggest usefulness of weighting genomic relationship matrix by using F_{ST}, especially when the trait is affected by a few numbers of QTL. The prediction accuracy of WssGBLUP methods is expected to increase by identifying and giving appropriate weight to QTL with major effects. Combining different test statistics into a single framework such as decomposition of multiple signals may help reduce false positives and pinpoint the QTL position with more precision.

KEY WORDS fixation index, genomic prediction, ssGBLUP, weighting.

INTRODUCTION

In a growing number of countries, genomic selection has become a routine method for predicting genomic breeding values (GEBVs) of selection candidates (Khansefid *et al.* 2020; Salek-Ardestani *et al.* 2021) due to its positive impact on genetic gain. However, different genomic prediction methods may have different predictive abilities, which are associated with factors such as the genetic architecture of traits and reference population size. Previous studies suggested that tracing selection footprints on the genome of two phenotypically divergent populations or sub-populations can help to detect genomic regions or QTLs related to traits that were under selection pressure (Chang *et al.* 2018; Chang *et al.* 2019). Specifically, signatures of selection tests can be employed for QTL mapping of oligogenic traits under selection (Walsh, 2021). Population fixation index (F_{ST}) as one of the most common cross-

population tests has been widely used to detect selection signatures in animals (Ghoreishifar *et al.* 2020a; Salek-Ardestani *et al.* 2020; Ghoreishifar *et al.* 2021). Recently, Chang *et al.* (2019) proposed a weighting genomic relationship matrix based on F_{ST} (hereinafter call WGBLUP$_{FST}$). In a simulated study, they demonstrated that WGBLUP$_{FST}$ could improve the accuracy of genomic prediction through using the F_{ST} values calculated from two subpopulations (i.e., top 5% and bottom 5% of individuals in the population based on their EBVs).

However, they did not compare their results with other weighting strategies. Additionally, the possible effect of heritability, reference population size, and distributions of QTL on predictive abilities were not considered in their study. Thus, this study aimed to investigate the performance of ssGBLUP and weighting genomic relationship matrix using F_{ST} and nonlinearA methods (hereinafter call WssGBLUP$_{NLA}$) in different scenarios (i.e., different heritability levels, reference population sizes and numbers of QTL).

MATERIALS AND METHODS

Simulation of populations

The QMSim (Sargolzaei and Schenkel, 2009) software used to simulate the genomic data. As shown in Table 1, data was generated in two steps. In the first step, a historical population with 2500 male and 2500 female animals was simulated and following 1000 generations of random mating, population size decreased linearly to 120 individuals to induce genome-wide linkage disequilibrium between SNP. Then, during 100 additional generations, the population size was expanded to 20000 animals, of which 19800 were females and 200 were males in the last historical generation. Mating pairs were random in the historical populations with non-overlapping generations, no selection and no migration. In the second step, of the animals in the last historical population, 12500 females and 30 males were randomly selected as founders to generate 15 overlapping generations (i.e., generations 1 to 15).

Each random mating of the selected parents with high EBV produced one progeny with 50% probability of being female or male. Sire and dam replacement rate were 0.2 and 0.5 per generation, respectively, and the effective population size was ~120.

Reference and validation populations

Generations 12-14 were included in the genomic prediction analyses. For these individuals (n=37500) phenotypes and pedigree records were also available. The animals in generation 14 (n=12500) and 1000 randomly selected animals from generation 15 were considered as reference and validation populations, respectively (Table 1).

Table 1 Simulation parameters of population structure and genomic data

Historical population (HP)	
Size of HP (number of generations)	5000 (0) 120 (1000) 20000 (1100)
Number of males (female) in the last Generation of HP	200 (19800)
Recent (founder) population	
Founder male (female) selected from HP	30 (12500)
Number of generations	15
Replacement rate for sire (dam)	0.5 (0.2)
Number of animals with pedigree (generation)	50000 (12-15)
Number of animals with phenotype (generation)	37500 (12-14)
Effective population size	~120
Number of replicates	5
Number of reference population (generation)	1500 (14) 5000 (14) 12500 (14)
Number of validation population (generation)	1000 (15)
Phenotypic variance	1.0
Simulated genome properties	
Total genome length (unit)	23.9 (Morgans)
Number of chromosomes	29
Chromosome's length	Similar to BovineSNP50k Chip
Number of markers per chromosome	Similar to BovineSNP50k Chip
Marker distribution	Evenly spaced
Total Number of QTL	10 50 500
Marker and QTL mutation rate (type)	2.5×10^{-5} (recurrent mutation)
QTL effects (shape parameter)	gamma distribution (0.4)

Three different numbers of individuals from generation 14 (1500, 5000 and 12500) were used in the reference population in different scenarios. The selection of animals to be in the reference and validation populations was at random but the same individuals were used in different ssGBLUP methods. For the validation population in all scenarios, the phenotypes were masked before genomic prediction analyses.

Genome and QTL simulation

The total length of the simulated genome was 23.19 Morgans which comprised of 29 chromosomes with equal length to the bovine autosomes (Lourenco *et al.* 2017). SNPs were uniformly distributed along the autosomes. The number of simulated SNPs was 54K, of which ~42.5K remained after quality control for minor allele frequency (MAF) < 0.05. Along with SNPs, bi-allelic QTL with MAF > 0.05 were randomly distributed along the simulated genome.

Six traits with different genetic architecture including different heritability (0.1 and 0.4) and numbers of QTL (10, 50 and 500) were simulated. QTL effects were sampled from a gamma distribution with a shape parameter of 0.4. Recurrent mutations for SNP and QTL were allowed with probability of 2.5×10^{-5}. The simulated phenotype for these traits comprised of the sum of an overall mean, the true breeding value (TBV) and a random residual. Each scenario was replicated five times.

Model and data analyses

The animal model below (Eq. 1) was used for genomic prediction:

$$y = Xb + Zu + e \qquad \text{Eq. 1}$$

Where:
y: n × 1 vector of observations.
b: n × 1 vector of fixed effects including overall mean.
u: q × 1 vector of random additive genetic effect driven from a normal distribution $u \sim N(0, H\sigma_u^2)$.
e: n × 1 vector of random residuals driven from a normal distribution $e \sim N(0, I\sigma_e^2)$, respectively.
X and Z: n × p and n × q design matrices which link the observations to fixed effects and random additive genetic effects, respectively.

Genomic prediction was performed using ssGBLUP and WssGBLUP. As described by Aguilar *et al.* (2010), in the mixed model equations for ssGBLUP, the pedigree-based relationship matrix (**A**) is replaced by a hybrid matrix called **H** matrix which allows to combine SNP and pedigree information**.** This matrix is constructed as follows:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau(\alpha G + \beta A_{22})^{-1} - \omega A_{22}^{-1} \end{bmatrix} \qquad \text{Eq. 2}$$

Where:
G^{-1}: inverse of genomic relationship matrix.
A^{-1} and A_{22}^{-1}: inverse of pedigree-based relationship matrix for all individuals in the pedigree and for the genotyped individuals, respectively.
τ and ω: scaling factors, both of which were set equal to one, as the default values in AIREMLF90 program.
α and β: blending factor to avoid singularity programs, which were set equal to 0.95 and 0.05, respectively.

In the equation above, G is constructed as follows:

$$G = MDM' / 2\Sigma p_j(1-p_j) \qquad \text{Eq. 3}$$

Where:
p_j: minor allele frequency for j^{th} marker.
M: allele frequency adjusted genotype matrix with elements including $0 - 2p_j$, $1 - 2p_j$ and $2 - 2p_j$ for genotypes AA, AB and BB, respectively.
D: diagonal matrix containing SNP weights with dimensions equal to the number of SNPs.

The ssGBLUP assumes equal variance for all SNPs, and therefore D is an identity (I) matrix.

WssGBLUP nonlinearA method

For WssGBLUP_{NLA}, SNPs weights were derived based on (VanRaden, 2008) formulae:

$d_{ii} = 1.125^{\frac{|\hat{u}_i|}{sd(\hat{u})}-2}$ using an iterative method proposed by Wang *et al.* (2012) as follows:

A) Set $D_t = I$ and $G_t = MDM' / 2\Sigma p_j(1-p_j)$
B) Compute GEBV using ssGBLUP approach
C) Compute SNP effects as $\hat{u} = \lambda DZ'G^{-1}\widehat{GEBV}$
D) Calculate SNP weights using NonlinearA: $d_{ii} = 1.125^{\frac{|\hat{u}_i|}{sd(\hat{u})}-2}$
E) Normalize D_{t+1}
F) Return back to step B with a new **G** matrix:
$G_{t+1} = MD_{(t+1)}M' / 2\Sigma p_j(1-p_j)$

WssGBLUP based on F_{ST}

For WssGBLUP_{FST}, SNP weights were derived based on population fixation index (F_{ST}) (Weir and Cockerham, 1984) as suggested by Chang *et al.* (2019). First, the breeding values (EBV) of the individuals in the reference population were estimated by BLUP animal model.
This model is similar to Eq. 1 but in the BLUP model, H was replaced by A, which is a pedigree numerator relationship matrix. Then, we assigned individuals in the reference population into three subpopulations including the bottom 5%, the middle 90% and the top 5% based on their EBV. The individuals with top and bottom 5% EBV were selected to calculate F_{ST} in PLINK v1.9 (Chang *et al.* 2015). Then, the F_{ST} values for all SNPs were scaled between 0 and 1 according to the maximum (F_{ST}^{max}) and minimum (F_{ST}^{min}) F_{ST} (Eq. 4), and then used as weights (D) to calculate weighted G (see Eq. 3).

$$Scaled\ F_{ST}^{i} = \frac{F_{ST}^{i} - F_{ST}^{min}}{F_{ST}^{max} - F_{ST}^{min}} \qquad \text{Eq. 4}$$

Accuracy and bias of genomic predictions

The correlation between GEBVs and TBVs of the validation animals were calculated and the average correlation over five replications (±SD) was reported as a measure of prediction accuracy. Additionally, the regression coefficients of TBVs on predicted GEBVs were calculated to

assess the dispersion bias of predictions. The regression coefficients were calculated using the "*lm*" R function and the average regression coefficient over five replications (±SD) was reported as a measure of bias of predictions in each scenario.

RESULTS AND DISCUSSION

The average prediction accuracy and SD in all scenarios including different number of QTL (10, 50 and 500) and different h^2 (0.1 and 0.4) within different reference population sizes (1500, 5000 and 12500) are shown in Figure 1.

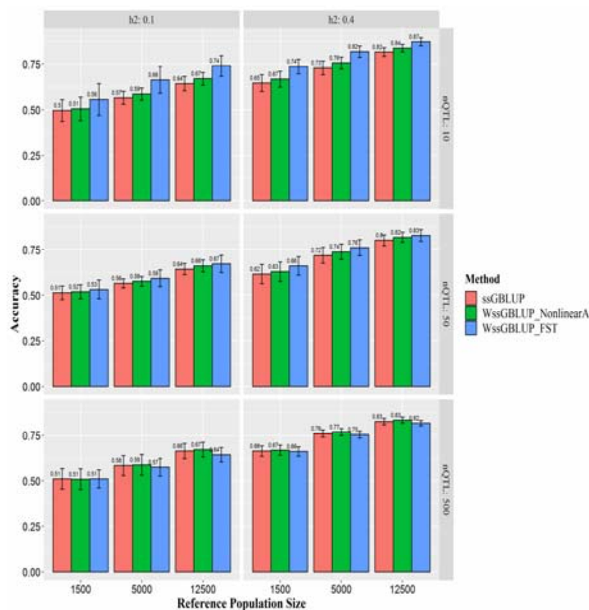

Figure 1 Accuracy of genomic predictions for different scenarios including varying numbers of QTL, reference population sizes and different heritability

In general, increasing the reference population size increased the accuracy of prediction in all scenarios regardless of prediction method. When the number of simulated QTL was low (i.e., 10 and 50 QTL), $\text{WssGBLUP}_{\text{FST}}$ outperformed $\text{WssGBLUP}_{\text{NLA}}$ and ssGBLUP. By increasing the size of reference population from 1500 to 12500 individuals, the prediction accuracy of $\text{WssGBLUP}_{\text{FST}}$ in the scenario with 10 QTL increased from 0.56 to 0.87 and in the scenario with 50 QTL increased from 0.53 to 0.83. For the same aforementioned reference population size, the prediction accuracy of $\text{WssGBLUP}_{\text{NLA}}$ increased from 0.51 to 0.84 in scenario with 10 QTL and increased from 0.52 to 0.82 in the scenario with 50 QTL. As expected, ssGBLUP method produced the least accurate predictions ranging from 0.50 to 0.82 and from 0.50 to 0.80 for 10 and 50 QTL, respectively. In the scenario of many QTL with small effects (i.e., 500 QTL), the $\text{WssGBLUP}_{\text{FST}}$ was no longer superior to $\text{WssGBLUP}_{\text{NLA}}$ and ssGBLUP. The prediction accuracy of $\text{WssGBLUP}_{\text{NLA}}$ and ssGBLUP were similar and in the range of 0.51 to 0.83.

In 500 QTL scenarios, the prediction accuracy obtained by $\text{WssGBLUP}_{\text{FST}}$ was slightly lower than other methods, and it was in the range of 0.51 to 0.82.

The calculated regression coefficients of TBVs on GEBVs are shown in Figure 2.

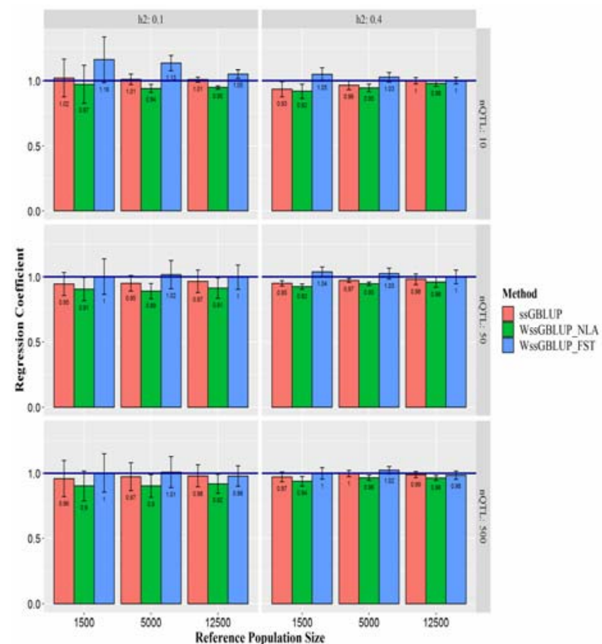

Figure 2 The regression coefficient of methods for different QTL scenarios, reference population sizes and heritability

A regression coefficient close to 1 means that GEBVs are not underestimated or overestimated. In general, all methods showed low prediction bias. In general, the regression coefficients for ssGBLUP and $\text{WssGBLUP}_{\text{NLA}}$ were slightly lower than 1, indicating that GEBVs were overestimated; and for $\text{WssGBLUP}_{\text{FST}}$ was higher than 1, indicating that GEBVs were underestimated. Increasing the number of QTL reduced the bias of predictions in $\text{WssGBLUP}_{\text{FST}}$ method for less heritable traits (h^2=0.1); while, had no or small effect on the bias of predictions for medium-to-high heritable traits. Moreover, increasing the size of reference population resulted in a reduction in bias of predictions in $\text{WssGBLUP}_{\text{FST}}$.

Giving different weights to SNPs to construct **G** has been reported to be useful to increase the accuracy of genomic prediction for traits with major QTL (Lourenco *et al.* 2017; Oget *et al.* 2019; Teissier *et al.* 2019; Mehrban *et al.* 2021). In this study, we investigated the accuracy and bias of genomic predictions using weighting methods called $\text{WssGBLUP}_{\text{NLA}}$ (VanRaden, 2008; Zhang *et al.* 2016) and $\text{WssGBLUP}_{\text{FST}}$ (Chang *et al.* 2019) for the traits with different genomic architecture and heritability, and different

reference population sizes. We also used regular ssGBLUP as the base prediction method. In our study, as expected, both $\text{WssGBLUP}_{\text{FST}}$ and $\text{WssGBLUP}_{\text{NLA}}$ outperformed ssGBLUP when the trait is controlled by a limited number of QTL (i.e., 10 and 50 QTL).

The results showed that the superiority of the $\text{WssGBLUP}_{\text{FST}}$ compared to $\text{WssGBLUP}_{\text{NLA}}$ depended on the genetic architecture of the trait, and size of the reference population. Result showed that, when a limited number of QTL were simulated, $\text{WssGBLUP}_{\text{FST}}$ produced more accurate GEBVs compared to $\text{WssGBLUP}_{\text{NLA}}$. In fact, by increasing the number of QTL from 10 to 50 QTL, $\text{WssGBLUP}_{\text{FST}}$ still outperformed $\text{WssGBLUP}_{\text{NLA}}$, but its superiority decreased from 7% to 1%. This could be explained by the QTL effects that were sampled from a gamma distribution where there are a small number of QTL with major effect explaining a larger proportion the genetic variance. Therefore, it seems that the F_{ST} outperforms in scenarios when the trait is governed by some major QTL. This, however, needs to be confirmed by real data. When 10 and 50 QTL were simulated, in general, the first major QTL explained around 40% and 15% of the total additive genetic variance, respectively (data not shown). To be more specific, in the Figure 3, F_{ST} values identified the first major QTL on chromosome 19 which explained about 85% of the total genetic variance of the trait.

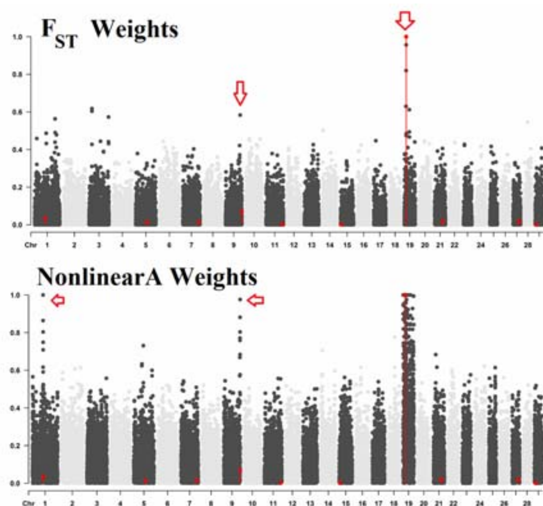

Figure 3 The Manhattan plot representing QTL effects (red circles) and SNP weights achieved by using NonlinearA and F_{ST} methods for 10-QTL scenario (h^2=0.1), and reference population size equal to 12500

In addition, F_{ST} identified the 2nd and 3rd major QTL on chromosomes 9 and 1, respectively. Regarding the fact that these QTL had large effects, their allele frequencies are expected to be so different in the top 5% and bottom 5% subpopulations of reference population. This could provide F_{ST} the power to identify these major QTL precisely (Ghoreishifar *et al.* 2021). Although $\text{WssGBLUP}_{\text{NLA}}$ identified the first 3 major QTL, and even QTL with smaller effects, e.g., a QTL on chromosome 5, it failed to fine map the signal on chromosome 19, and gave almost equal weights to the three major identified QTL. Therefore, it seemed that F_{ST} could identify major QTL more precisely. This might explain why the prediction accuracy of $\text{WssGBLUP}_{\text{NLA}}$ was less than that of $\text{WssGBLUP}_{\text{FST}}$ when a limited number of QTL were simulated, and even explain that why the superiority of $\text{WssGBLUP}_{\text{FST}}$ to $\text{WssGBLUP}_{\text{NLA}}$ decreases from 7% to 1% when the number of simulated QTL increased from 10 to 50 QTL. Note that the weights of F_{ST} and nonlinearA as well as QTL effects were scaled between zero and one in Figure 3.

It should also be accounted that despite using 1250 individuals (i.e., 5% top and 5% bottom of 12500 individuals in the reference population) for F_{ST} calculation, false positive signals are more likely to be introduced to the prediction model (i.e., chromosomes 1 and 3 in Figure 3) that might reduce the prediction accuracy and increase the bias of GEBVs. To deal with this challenge, and in order to reduce the number of detected false positive QTL signals, application of different selection signature methods and combining them into a framework called DCMS (de-correlated composite of multiple signals) (Ma *et al.* 2015) might be an option which can be further studied in the future. Ma *et al.* (2015) reported that the resolution of selection signature mapping and the power of detecting selection signals were improved by using DCMS compared to most single statistics, such as F_{ST}. Ghoreishifar *et al.* (2020b) reported that incorporating *p*-values of different statistics in a single DCMS framework may help select and prioritize candidate genes. Moreover, composite measures such as DCMS have been reported to identify the causal variants (i.e., the variants under selection in the detected signature regions) more precisely. It was also reported that by increasing the marker density, the power of DCMS method could be increased. Generally, F_{ST} could be used to identify variants that are fixed or close to fixation. Therefore, using other methods such as iHS and xpEHH methods and combining them into a DCMS framework could help to detect QTL with intermediate frequency (Ma *et al.* 2015) as well. In general, the prediction accuracy of WssGBLUP methods is expected to be increased by identifying and giving appropriate weight to QTL with major effects in addition to reducing false positive rate in QTL mapping.

Based on simulation, three different number of QTL representing oligogenic traits affected by small number of QTL (i.e., 10 and 50 QTL) and polygenic traits affected by many QTL with small effects (500 QTL). Some studies simulated over 5,000 QTL to mimic complex traits. However, we did not simulate more than 500 QTL because we

observed that by increasing the number of QTL from 10 to 500, the superiority of weighting methods to ssGBLUP decreased ($WssGBLUP_{FST}$) or remained constant ($WssGBLUP_{NLA}$). Hence, weighting ssGBLUP is not recommended for polygenic traits unless the QTL were detected, and their weights could be calculated precisely. Given the limitations in detecting QTL with small effects, for the traits controlled by the number of QTL greater than 500, it is unreasonable to use $WssGBLUP_{FST}$. It is worth noting that selection for polygenic traits would leave only minor footprints because of the selection for numerous regions with lower intensity across the genome (Kemper *et al*. 2014; Ghoreishifar *et al*. 2020b). As a result, identification of these QTLs with small effects is difficult to track with F_{ST}.

CONCLUSION

The F_{ST} could be a powerful method to detect major QTL compared to nonlinearA method, while the latter could be more useful to identify QTL with smaller effects. This could be attributed to superiority of F_{ST} over NonlinearA for genomic predictions of the traits explained by a few QTLs. The false positive QTL signals, undetected QTLs and inaccurate weights are potentially restricting the usefulness of WssGBLUP for genomic predictions of oligogenic traits. Therefore, identification of major QTL by using high-density markers and application of multiple methods such as different selection signature statistics and even combining them with NonlinearA might help to detect QTL and consequently improve the genomic prediction for oligogenic traits in WssGBLUP.

ACKNOWLEDGEMENT

We have received no special funding to conduct this study.

REFERENCES

Aguilar I., Misztal I., Johnson D., Legarra A., Tsuruta S. and Lawlor T. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* **93(2),** 743-752.

Chang C.C., Chow C.C., Tellier L.C., Vattikuti S., Purcell S.M. and Lee J.J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*. **4(1),** 7-15.

Chang L.Y., Toghiani S., Hay E.H., Aggrey S.E. and Rekaya R. (2019). A weighted genomic relationship matrix based on Fixation Index (Fst) prioritized SNPs for genomic selection. *Genes*. **10(11),** 922-931.

Chang L.Y., Toghiani S., Ling A., Aggrey S.E. and Rekaya R. (2018). High density marker panels, SNPs prioritizing and accuracy of genomic selection. *BMC Genet.* **19(1),** 1-10.

Ghoreishifar S.M., Eriksson S., Johansson A.M., Khansefid M., Moghaddaszadeh-Ahrabi S., Parna N., Davoudi P. and Javanmard A. (2020a). Signatures of selection reveal candidate genes involved in economic traits and cold acclimation in five Swedish cattle breeds. *Genet. Sel. Evol.* **52(1),** 1-15.

Ghoreishifar S.M., Moradi-Shahrbabak H., Fallahi M.H., Jalil Sarghale A., Moradi-Shahrbabak M., Abdollahi-Arpanahi R. and Khansefid M. (2020b). Genomic measures of inbreeding coefficients and genome-wide scan for runs of homozygosity islands in Iranian river buffalo, Bubalus bubalis. *BMC Genet.* **21(1),** 16-22.

Ghoreishifar S.M., Rochus C.M., Moghaddaszadeh-Ahrabi S., Davoudi P., Ardestani S.S., Zinovieva N.A., Deniskova T.E. and Johansson A.M. (2021). Shared ancestry and signatures of recent selection in gotland sheep. *Genes*. **12(3),** 433-442.

Kemper K.E., Saxton S.J., Bolormaa S., Hayes B.J. and Goddard M.E. (2014). Selection for complex traits leaves little or no classic signatures of selection. *BMC Genom.* **15(1),** 246-254.

Khansefid M., Goddard M.E., Haile-Mariam M., Konstantinov K.V., Schrooten C., de Jong G., Jewell E.G. O'Connor E. Pryce J.E. Daetwyler H.D. and MacLeod I.M. (2020). Improving Genomic Prediction of Crossbred and Purebred Dairy Cattle. *Fron. Genet.* **11,** 1-14.

Lourenco D., Fragomeni B., Bradford H., Menezes I., Ferraz J., Aguilar I., Lourenco S. and Misztal I. (2017). Implications of SNP weighting on single‐step genomic predictions for different reference population sizes. *J. Anim. Breed. Genet.* **134(6),** 463-471.

Ma Y., Ding X., Qanbari S., Weigend S., Zhang Q. and Simianer H. (2015). Properties of different selection signature statistics and a new strategy for combining them. *Heredity*. **115(5),** 426-436.

Mehrban H., Naserkheil M., Lee D.H., Cho C., Choi T., Park M. and Ibáñez-Escriche N. (2021). Genomic prediction using alternative strategies of weighted single-step genomic BLUP for yearling weight and carcass traits in Hanwoo beef cattle. *Genes.* **12(2),** 266-275.

Oget C., Teissier M., Astruc J.M., Tosser-Klopp G. and Rupp R. (2019). Alternative methods improve the accuracy of genomic prediction using information from a causal point mutation in a dairy sheep model. *BMC Genom.* **20(1),** 1-14.

Salek Ardestani S., Aminafshar M., Zandi Baghche M., Banabazi M.H., Sargolzaei M. and Miar Y. (2020b). Whole-genome signatures of selection in sport horses revealed selection footprints related to musculoskeletal system development processes. *Animals*. **10(1),** 53-64.

Salek Ardestani S., Jafarikia M., Sargolzaei M., Sullivan B. and Miar Y. (2021). Genomic prediction of average daily gain, back-fat thickness, and loin muscle depth using different genomic tools in Canadian Swine populations. *Front. Genet.* **12,** 735-743.

Sargolzaei M. and Schenkel F.S. (2009). QMSim: A large-scale

genome simulator for livestock. *Bioinformatics*. **25(5),** 680-681.

Teissier M., Larroque H. and Robert-Granie C. (2019). Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in French dairy goats. *J. Dairy Sci*. **102(4),** 3142-3154.

VanRaden P.M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci*. **91(11),** 4414-4423.

Walsh J.B. (2021). Genomic selection signatures and animal breeding. *J. Anim. Breed. Genet.* **138,** 1-3.

Wang H., Misztal I., Aguilar I., Legarra A. and Muir W. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* **94(2),** 73-83.

Weir B.S. and Cockerham C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*. **38(6),** 1358-1370.

Zhang X., Lourenco D., Aguilar I., Legarra A. and Misztal I. (2016). Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. *Front. Genet.* **7,** 151-162.