Genomic Prediction of Additive and Dominant Effects on Wool and Blood Traits in Alpine Merino Sheep

Shaohua Zhu [1,2], Hongchang Zhao [2], Mei Han [2], Chao Yuan [2], Tingting Guo [2], Jianbin Liu [2], Yaojing Yue [2], Guoyan Qiao [2], Tianxiang Wang [3], Fanwen Li [3], Shuangbao Gun [1]* and Bohui Yang [2]*

[1] *College of Animal Science and Technology, Gansu Agricultural University, Lanzhou, China,* [2] *Animal Science Department, Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou, China,* [3] *Gansu Provincial Sheep Breeding Technology Extension Station, Sunan, China*

Dominant genetic effects may provide a critical contribution to the total genetic variation of quantitative and complex traits. However, investigations of genome-wide markers to study the genomic prediction (GP) and genetic mechanisms of complex traits generally ignore dominant genetic effects. The increasing availability of genomic datasets and the potential benefits of the inclusion of non-additive genetic effects in GP have recently renewed attention to incorporation of these effects in genomic prediction models. In the present study, data from 498 genotyped Alpine Merino sheep were adopted to estimate the additive and dominant genetic effects of 9 wool and blood traits via two linear models: (1) an additive effect model (MAG) and (2) a model that included both additive and dominant genetic effects (MADG). Moreover, a method of 5-fold cross validation was used to evaluate the capability of GP in the two different models. The results of variance component estimates for each trait suggested that for fleece extension rate (73%), red blood cell count (28%), and hematocrit (25%), a large component of phenotypic variation was explained by dominant genetic effects. The results of cross validation demonstrated that the MADG model, comprising additive and dominant genetic effects, did not display an apparent advantage over the MAG model that included only additive genetic effects, i.e., the model that included dominant genetic effects did not improve the capability for prediction of the genomic model. Consequently, inclusion of dominant effects in the GP model may not be beneficial for wool and blood traits in the population of Alpine Merino sheep.

Keywords: Alpine Merino sheep, additive effects, dominant effects, prediction accuracy, genomic prediction

INTRODUCTION

In classical models of quantitative or complex trait genetics, the phenotypic value of each trait is controlled by a large number of loci; moreover, the interaction and alternative splicing of genes also play an extremely essential role (1). The phenotypic value is also affected by non-genetic and environmental factors (2). Therefore, it is not possible to select the top-quality animal and plant population by genotype on a single marker (or gene) or through simply observing their extrinsic features. The selection is based on the predicted total effect of the loci within an individual or their

estimated breeding values (BVs) (2). Genomic selection (GS) adopts markers covering the entire genome, so that these markers can be used to explain all genetic variations (3). Compared with traditional selection methods, it has higher prediction accuracy; in addition, it could reduce the generation interval and increase the genetic progress (4, 5).

GS has improved modern breeding programs and has made a great contribution to improving the accuracy of BVs (6, 7), especially for the early selection of young individual domesticated animals (8). Genetic effects including additive effects and non-additive effects are both important for analyzing the genetic mechanism of livestock complex traits through whole-genome markers or for GP (9). However, in production and application, more attention is generally paid to additive effects, because additive genetic effects reflect the breeding value of individuals (10). Although the studies on non-additive effects has made progress in dairy cattle and beef cattle (11, 12), due to incomplete pedigrees or unreliable records, restricted calculations, and other problems, few studies have focused on the genetic, non-additive effects for sheep populations (13). The dominant effect represents interactions at the same locus; the epistatic effects represent interactions between different loci. Previous studies have suggested that non-additive genetic effects could provide an essential influence on the total variation of complex traits (14–16).

With the publication of the whole-genome sequence of dairy cows and the continuous upgrading of commercial SNP microarrays, GP and GS have been adopted on a large scale in conventional breeding programs of dairy cows (17, 18). This has also promoted the development of SNP microarrays for a variety of other livestock. To date, the associated GS technology has been gradually expanded to livestock such as beef cattle (19), pigs (20), goats (21), and sheep (22), but their development still lags far behind that of dairy cattle. The Alpine Merino sheep, the subject of the present study, possesses a mixture of Australian Merino and Tibetan sheep ancestry. They quickly adapted to the cold Qinghai-Tibet Plateau and have subsisted in high-altitude and hypoxic environments for generations. However, since the technology of GS has not been popularized in sheep, the Alpine Merino sheep population is still during the traditional phenotypic selection period. Therefore, genomic analysis to identify wool traits associated with cold tolerance and erythrocyte traits associated with adaptation to high-altitude hypoxia, critical for the selection and breeding of this population, has been used. To the best of our knowledge, GP or GS studies of Alpine Merino sheep have not so far been reported. With the availability of SNP genotypes, it is possible to explore the additive and dominant genetic effects of marker loci and thus estimate the genetic effects of each marker (23). Furthermore, in a manner similar to the construction of a genome additive relationship matrix, a genomic dominant effect matrix could also be integrated into a genomic best linear unbiased prediction (GBLUP) model for genomic analysis research (24, 25).

For this study, components of the genetic variance of the Alpine Merino sheep dataset were estimated, including additive and dominant genetic variance, then the accuracy of the BVs estimated by two models compared using cross validation. The main objective of this study was to explore the impact of additive and dominant genetic effects on the accuracy of GP and the optimization methods of GP for the Alpine Merino sheep population.

MATERIALS AND METHODS

Ethics Statement

All animal work conducted in the present study was performed in accordance with the guidelines for the care and use of laboratory animals promulgated by the State Council of the People's Republic of China. This research was approved by the Animal Management and Ethics Committee of Lanzhou Institute of Animal Husbandry and Veterinary Medicine, Chinese Academy of Agricultural Sciences (license number: 2019-008).

Phenotypic Data Collection

The initial phenotypic dataset was derived from animals on the HuangCheng pasture in Gansu Province, China, part of the Sheep Breeding Technology Extension Station. The dataset consisted of 11,200 individuals and included their dates of birth, region (defined by herd), and sex. In the present study, the phenotypes of 498 offspring for 26 breeding rams were collected from 7 different herds, consisting of 295 rams and 203 ewes. Most of them were half-sibs, and very few (<12%) were full-sibs, the age of which were around 14 months. A blood sample (5 mL) from each sheep was collected from the jugular vein and immediately transferred to a vacuum blood collection tube (Yuli Medical Equipment Company Ltd., Jiangsu Province, China). A standard set of red blood cell data was recorded using an H-100IV diff whole blood analyzer (Sysmex, Kobe, Japan) within 24 h of sample collection, after which the remaining blood sample was stored at −20°C (26). Four erythrocyte parameters closely associated with adaptation to high-altitude hypoxia were selected as blood trait phenotypic data, including red blood cell count (RBC), hematocrit (HCT), mean cell hemoglobin (MCH), and mean cell hemoglobin concentration (MCHC). Every year in July, when the lambs grow to 14 months of age, they reach the period of shearing. According to the agricultural industry standard of the People's Republic of China (No. NY/T 1236-2006), a 150–200-g sample of wool was collected from the abdomen of each sheep, weighed, and stored in Ziplock bags. Within 5 days of sample collection, they were sent to the National Animal and Rural Ministry of Animal and Fur Quality Supervision and Inspection Center (Lanzhou, China) for weighing, cleaning, and quality testing. Five parameters were selected to represent the phenotypic data of wool traits, including staple length (SL), clean fleece weight rate (CFWR), mean fiber diameter (FD), fleece breaking strength (FBS), and fleece extension rate (FER). **Supplementary Table 1** shows the detailed definition of the above wool traits parameters.

Genotyping and Quality Control

In total, 498 Alpine Merino sheep were genotyped using a custom Affymetrix HD 630K microarray. The genotyping platform adopted for analysis was based on the GeneTitan System (Santa Clara, CA, USA) Array Plate Processing Workflow

from Thermo Fisher (Affymetrix). Prior to statistical analysis, SNPs were pre-processed using PLINK v1.9b4 software (27). Samples were eliminated where SNPs had a call rate lower than 95%, a Hardy–Weinberg equilibrium $P < 10e-6$, a minor allele frequency <0.01, and those with more than 10% of their genotype missing. Moreover, in order to eliminate the potential biased analysis caused by gender and maternal effects, the X, Y, and mitochondrial chromosomes were excluded from analysis (11). Following filtering and quality control, a total of 498 individual animals with 441,740 autosomal SNPs were retained.

Statistical Methods for GP

In this study, we explored the possibility of combining both additive and dominant genetic effects in the genomic evaluation and further compared the accuracy of GP via different models. Nine traits of 498 samples were used in (1) estimating the variance of components in the genotype dataset of Alpine Merino sheep, including additive and dominant variance, and (2) analyzing the accuracy of GP using 5-fold cross validation and then comparing the accuracy of estimation of BVs between the prediction model using additive effects and that when combining both additive and dominant effects. Both models were evaluated within a GBLUP framework. Replicate measurements were not available for the sheep; thus, permanent environmental effects were not modeled. Samples were of different genders and from different herds, factors that altered phenotypes in a fixed manner, so system environmental effects were added to the framework.

The components of variance of additive genetic effect (σ_A^2) and residual effect (σ_E^2) were estimated adopting the MAG model.

$$y = Xb + Zu + e \tag{1}$$

where y represents the phenotypic value of the individual and b refers to the vector of fixed effects, since the individuals involved in the current study were from different herds; the fixed effects include the gender (male and female) of each individual and the different herds (H1–H7). u is the vector of breeding values of the individual, e is the vector of residual effects, X is the design matrix corresponding to fixed effects, and Z is the design matrix corresponding to breeding values. Covariance matrices of additive effects were $Var(u) = G\sigma_a^2$, where G is the matrix of the genomic relationship, calculated based on the approach of VanRaden (28) and using R package "HIBLUP" (https://github.com/xiaolei-lab/hiblup):

$$G = \frac{W_1 W_1^T}{2 \sum_{i=1}^m p_i (1 - p_i)} \tag{2}$$

where W_1 is the matrix of additive genetic effect markers, with dimensions of the number of individuals (n) by the number of loci (m), and p_i is the minor allele frequency (MAF) value of locus i.

In addition, the components of variance of the additive genetic effect (σ_A^2), dominant genetic effects (σ_D^2), and residual effect (σ_E^2) were estimated using the MADG model:

$$y = Xb + Zu + Zv + e \tag{3}$$

As with the MAG model, in Equation (3), y, b, u, e, X, and Z represent the same parameters as those defined in Equation (1), while v refers to the vector of the dominant effect of an individual. The covariance matrices of dominant effects were $Var(v) = D\sigma_d^2$, where D represents the matrix of the genomic dominant relationship (25), also calculated using R package "HIBLUP" (https://github.com/xiaolei-lab/hiblup):

$$D = \frac{W_2 W_2^T}{4 \sum_{i=1}^m p_i^2 (1 - p_i)^2} \tag{4}$$

where W_2 represents the matrix of dominant genetic effect markers and where, n, m, and p_i are the same as those defined in Equation (2).

Accuracy of Breeding Value Predictions by Cross-Validation

The accuracy of GP was evaluated by 5-fold cross validation. The dataset was randomly divided into 5 approximately equally sized subgroups (each subgroup contained about 100 individuals), for each cross validation, of which 4 were considered as training groups (reference population) to estimate parameters, and the remaining group (candidate population) was used to validate samples. The population was generally randomly divided when performing 5-fold cross validation. A number of validation samples had offspring in the training population, causing in that case the cross validation to be based on offspring, which would exaggerate the accuracy of the prediction (29). In the current population, individuals were collected from the same generation, ensuring that verification samples did not have offspring in the training group. In this method of grouping, the 5-fold cross validation evaluated the accuracy of ewes and rams of the same generation. Therefore, it was effective in limiting the bias in accuracy caused by offspring in the training group (11).

Based on the MAG and MADG models, the breeding values of the validation group were predicted and their respective components of variance estimated. In addition, we performed the 5-fold cross validation described above twice to ensure randomness of the verification group. Finally, 10 different values of accuracy were calculated for each trait, the mean value recorded as the final accuracy.

RESULTS

Phenotypic Statistics and Genotypic Characteristics

Table 1 displays descriptive statistics of the phenotypic measurements of the sheep, including mean of each parameter, abbreviations of each trait standard error, coefficient of variation, and number of individuals. In detail, it shows five wool traits and four blood traits. For wool traits, the coefficient of variation (CV) ranged from 0.1 (staple length) to 0.3 (Fleece extension rate), and the standard error (SE) ranged from 0.1 (mean fiber diameter) to 0.4 (fleece breaking strength); for blood traits, the CV ranged from 19.9 (mean cell hemoglobin) to 29.9 (hematocrit), and the SE is ranged from 0.0 (hematocrit) to 4.7 (mean cell hemoglobin concentration).

Prediction of Breeding Values and Total Genetic Values

Estimated components of variance for the MAG and MADG models are presented in **Table 2**. Additive variance (showed as a proportion of total genetic variance) for wool traits ranging from 6.88% (fleece extension rate) to 49.5% (staple length) and for blood traits from 13.31% (red blood cell count) to 27.94% (mean cell hemoglobin concentration). Dominant variance (expressed as a percentage of total genetic variance) for wool traits ranged from 0.00 to 73.46%. The dominant variance of staple length and mean fiber diameter were extremely low, approximately equal to 0.00%. The dominant variance of clean fleece weight rate, fleece breaking strength, and fleece extension rate were 4.22, 10.06, and 73.41%, respectively. Dominant variance for blood traits ranged from 0.00 to 27.99%. The dominant variance of mean cell hemoglobin and mean cell hemoglobin concentration were approximately equal to 0.00%. The dominant variance of red blood cell count and hematocrit were 27.99 and 24.99%, respectively.

Table 3 compares the capability of GP for the two models. The accuracy calculated by 5-fold cross validation varied with the different traits. For wool traits, the prediction accuracy of FER was least (MAG = 0.03, MADG = 0.01) while that of SL was greatest (MAG = 0.25, MADG = 0.25). For blood traits, the prediction accuracy for RBC was least (MAG = 0.04, MADG = 0.02), while that of MCH was greatest (MAG = 0.16, MADG = 0.15). Of all 9 traits, including wool and blood traits, FER and RBC displayed relatively low prediction accuracy (<0.1), while the others displayed low to medium accuracy (0.11–0.25). In addition, the MAG model exhibited accuracy slightly higher than that of the MADG model in all traits (**Figure 1**), regardless of whether heritability was high or low.

TABLE 1 | Descriptive statistics of phenotypic values of traits.

Trait	Abbreviation	Mean ± SD[a]	CV[b] (%)	SE[c]	Number
Staple length (mm)	SL	83.5 ± 9.4	0.1	0.4	494
Clean fleece weight rate (%)	CFWR	66.1 ± 6.4	0.1	0.3	494
Mean fiber diameter (mm)	FD	21.4 ± 2.2	0.1	0.1	491
Fleece breaking strength (N/ktex)	FBS	33.3 ± 8.0	0.2	0.4	491
Fleece extension rate (%)	FER	19.5 ± 5.1	0.3	0.2	493
Red blood cell count (10^{12}/L)	RBC	7.7 ± 1.7	22.1	0.1	496
Hematocrit (%)	HCT	0.3 ± 0.1	29.9	0.0	492
Mean cell hemoglobin (pg)	MCH	13.2 ± 2.6	19.9	0.1	494
Mean cell hemoglobin concentration (g/L)	MCHC	377.5 ± 103.8	27.5	4.7	489

[a]SD, standard deviation.
[b]SE, standard error.
[c]CV, coefficient of variation.

TABLE 3 | Comparison of prediction accuracies of 9 traits.

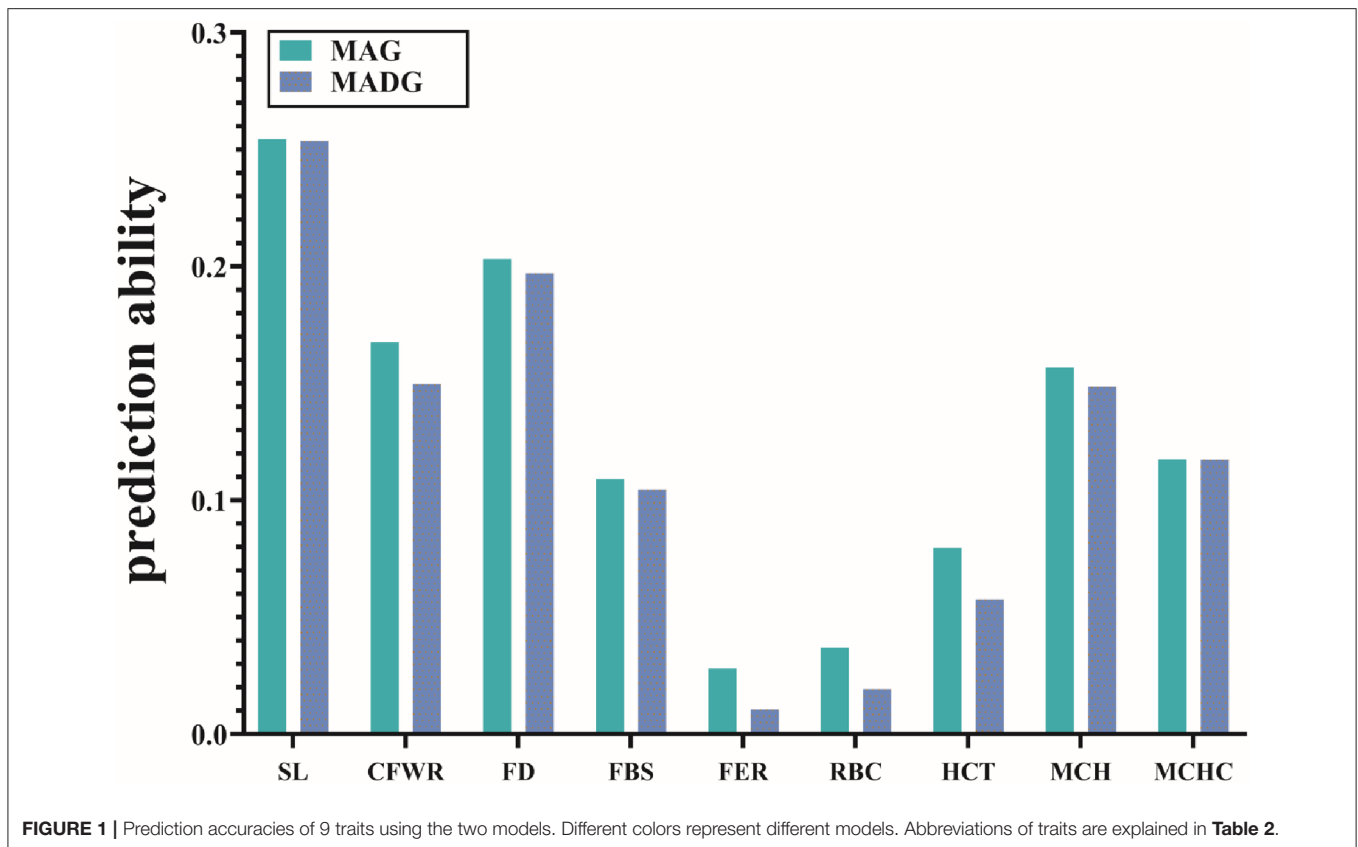| Trait[a]	Prediction accuracy[b]	
	MAG	**MADG**
SL	0.25 (0.02)	0.25 (0.02)
CFWR	0.17 (0.03)	0.15 (0.03)
FD	0.20 (0.02)	0.20 (0.02)
FBS	0.11 (0.03)	0.10 (0.04)
FER	0.03 (0.02)	0.01 (0.03)
RBC	0.04 (0.02)	0.02 (0.04)
HCT	0.08 (0.02)	0.06 (0.03)
MCH	0.16 (0.03)	0.15 (0.03)
MCHC	0.12 (0.02)	0.12 (0.02)

[a]Abbreviations of traits explained in **Table 2**.
[b]SE are in parenthesis.

TABLE 2 | Estimates of additive and dominant components of variance obtained using HIBLUP for MAG and MADG models.

Trait[a]	MAG			MADG			
	σ_a^2(SE)	σ_e^2(SE)	$\frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2}$ (SE)[b]	σ_a^2(SE)	σ_d^2(SE)	σ_e^2(SE)	$\frac{\sigma_a^2}{\sigma_a^2+\sigma_d^2+\sigma_e^2}$ (SE)[b]
SL	40.96 (11.16)	41.79 (9.78)	0.50 (0.12)	30.25 (17.93)	0.00 (34.61)	48.81 (21.45)	0.00 (0.44)
CFWR	11.93 (4.68)	17.79 (4.24)	0.40 (0.15)	11.41 (8.54)	1.25 (17.53)	17.05 (11.27)	0.04 (0.59)
FD	1.54 (0.55)	2.61 (0.51)	0.37 (0.13)	4.11 (3.02)	0.00 (5.21)	4.43 (3.16)	0.00 (0.61)
FBS	18.03 (7.99)	35.34 (7.45)	0.34 (0.14)	15.50 (14.46)	5.36 (28.20)	32.44 (17.79)	0.10 (0.53)
FER	1.34 (1.96)	18.10 (2.22)	0.07 (0.10)	0.00 (8.11)	22.29 (12.24)	8.07 (6.98)	0.73 (0.40)
RBC	0.35 (0.31)	2.25 (0.33)	0.13 (0.12)	0.06 (0.71)	0.73 (1.67)	1.81 (1.07)	0.28 (0.64)
HCT	0.001 (0.001)	0.005 (0.008)	0.24 (0.12)	0.001 (0.001)	0.00 (0.00)	0.004 (0.002)	0.25 (0.48)
MCH	1.34 (0.78)	4.48 (0.78)	0.23 (0.13)	2.03 (2.61)	0.00 (6.10)	5.59 (3.97)	0.00 (0.81)
MCHC	2867.17 (1420.81)	7395.22 (1371.06)	0.28 (0.13)	4388.68 (4504.31)	0.00 (1001.76)	8854.82 (1428.41)	0.00 (0.76)

[a]SL, staple length; CFWR, clean fleece weight rate; FD, mean fiber diameter; FBS, fleece breaking strength; FER, fleece extension rate; RBC, red blood cell count; HCT, hematocrit; MCH, mean cell hemoglobin; MCHC, mean cell hemoglobin concentration.
[b]Heritability, the ratio of the additive effect variance to the total phenotypic variance.

FIGURE 1 | Prediction accuracies of 9 traits using the two models. Different colors represent different models. Abbreviations of traits are explained in **Table 2**.

DISCUSSION

Genome Relationship Matrix

Genome-wide dense markers have replaced pedigree information and have provided a new methodology for estimating additive and dominant genetic effects, anticipated to improve the accuracy of genomic predictions (30). The technique of construction of an additive genomic relationship matrix by whole-genome markers (28, 31) then using a linear mixed model to estimate individual breeding values has been widely used in genomic prediction and selection (32–34). In the present study, in addition to an additive genomic relationship matrix, we also constructed a dominant relationship matrix via genome-wide SNP markers. The genome relationship matrix represents confirmed gene sharing, rather than merely conceptualized or predicted ancestral sharing (24). In contrast to pedigree-based individual relationship matrices, a genomic relationship matrix based on genome-wide dense markers is able to capture genetic links from unknown common ancestors (24), not available in an individual relationship matrix based on pedigree. Furthermore, a genomic relationship matrix is suitable for not only populations with pedigree information but also those without pedigree information. It is particularly useful for the study of livestock populations and even wild populations that lack or have inaccurate pedigree information (35, 36). The technique for construction of an epistatic genome relation matrix allows for only approximate

calculations when considering a large number of markers. Where epistasis between several specific markers is modeled using a required technique, an epistatic genomic relationship could only be approximated, as it is difficult to construct a precise epistatic relationship matrix (37, 38). Therefore, an epistatic relationship matrix was not constructed in the current study.

The use of an additive and dominant genetic relationship matrix in a typical linear mixed model (such as the GBLUP model) is capable of estimating additive and dominant components of variance. Goddard and Hayes et al. established that the GBLUP model is equivalent to a linear random regression model, assuming that each SNP causes an effect, and the effects of all SNP effects follow a normal distribution with equal variances (39, 40). The predictability of the GBLUP model may be unsatisfactory in cases where a few markers have large effects and the majority have little effect or are even ineffective. However, the study of actual data from German Holstein cows has demonstrated that the GBLUP model, which directly estimates genomic breeding value, was able to fully utilize the related information (41). Thus, the GBLUP model has comparative advantages in the case of close relationships between reference and candidate populations (42, 43). Considering that the population involved in the current study was from the same breed, the method proposed in the present study may not be the only one,

but it was an appropriate approach for estimation of additive and dominant variances and for prediction of genomic breeding values.

Additive and dominant variance components were estimated for 9 traits of interest in the Alpine Merino sheep population, including 5 wool traits and 4 blood traits. Due to complex genetic mechanisms and differences between species, to date, few studies have estimated genetic variance components in sheep blood traits. Safari and Fogarty et al. collected and summarized a number of studies on the genetic parameters of important traits in sheep. Their reported results suggested that the weighted average heritability of FD (0.51–0.59) was based on 41 and 43 estimates, respectively; the weighted average heritability of SL (0.46–0.48) was based on 21 estimates; the weighted average heritability of FBS (0.34) was based on 11 estimates; and the weighted average heritability of CFWR (0.34–0.51) was based on 36 and 43 estimates, respectively (44, 45). In this study, the heritability of 4 wool traits (**Table 2**) was close to the results reported in previously published literature except that 0.37 for FD was slightly lower. In particular, SL (0.50) and FBS (0.34) were very close to the results reported in the literature. The results suggested that although the dataset was only a small proportion of the entire Alpine Merino population, heritability estimates were still reliable. For the 9 traits involved in the current study, the estimated values of dominant variance ranged from 0 to 73.41% of the total genetic variation. It has been reported in the literature that estimated variation in dominant milk production traits in Holstein cows ranged from 1.4 to 42.9% of the total genetic variance, with a wide range of variation (46–48). This is consistent with the research results in the present study, but with a smaller range. Two reasons may explain why the range of our results was larger than that reported in the literature. (1) An excellent rearing bull is able to participate in the breeding of a large number of cows through semen cryopreservation technology, hard to achieve in the Alpine Merino sheep population. In a similar-sized population, the male-to-female mating ratio of Alpine Merino sheep was considerably larger than that of Holstein cows, proving richer genetic diversity to the Alpine Merino sheep, leading to greater population heterozygosity and QTL alleles with greater intermediate frequencies (49). (2) Research reported in the literature was based on an individual relationship matrix constructed using a pedigree dataset, then a conventional BLUP method used to estimate the various components of variance, while the individual relationship matrix was constructed from genomic information in the present study, which is expected to improve the ability to estimate components of variance, including additive and dominant variance, and also effectively reduce potential confusion about additive effects and residuals. This confusion can also lead to the different estimations of components of variance (11). Previous studies have shown a significant contribution of non-additive genetic variation. It has been reported that in a Duroc pig population, the non-additive variance of the majority of meat quality and carcass traits accounts for more than 50% of total QTL variance (50). The ratios of dominant variance

to additive genetic variance ranged from 15% (21-day litter weight) to 57% (interval between parities) for reproductive and growth traits in South African Duroc pigs (16). In chickens, QTL analysis demonstrated that non-additive genetic effects explained greater variation in those younger than 46 days, while additive genetic effects explained the principal proportion of later life genetic variation (51, 52). These results indicate that non-additive genetic variation is extremely important in complex traits.

GP Results and Accuracy of Prediction

To the best of our knowledge, studies of GP or GS about high-altitude hypoxia adaptation in sheep were rarely reported, so we did not find reference statistics related to sheep blood traits. However, for wool traits, the statistical averages of FD and FBS calculated by Daetwyler et al. in their study on genetic parameter estimation of Australian Merino sheep were 17.3 and 33.4, respectively (53). The statistical averages of SL and FBS obtained by Moghaddar et al. in the genomic prediction study of wool traits of Poll Dorset (PD) and Merino and White Suffolk (WS) were 80.93–98.57 and 33.80–35.61, respectively (54); in addition, according to Hamadani and coworkers in the study of Rambouillet sheep heritability estimation, the statistical average of FD is 21.26 (55). In the current study, the statistical results (**Table 1**) were consistent with the values calculated from the above studies, and it suggested that the statistical values of the phenotypic measurements were reliable.

In the present study, the accuracy of GP was evaluated by 5-fold cross validation, the results of which are presented in **Table 3**. The results of the GP study on wool traits of different breeds of sheep from Daetwyler et al. suggested that for FD traits, the accuracy of prediction ranged from 0.23 to 0.79, and for FBS traits, it ranged from−0.01 to 0.43 (53). Moreover, Moghaddar and his colleagues estimated the BVs of wool trait such as FD, SL, and FBS of Merino, Border Leicester (BL), and WS; the prediction accuracy of FD, SL, and FBS was between 0.39 and 0.50, 0.25 and 0.61, and 0.10 and 0.49, respectively (54). In the current study, the prediction accuracy of FD, SL, and FBS was 0.20, 0.25, and 0.11 respectively, which were close to the results reported in the previous literatures. For these 9 traits, regardless of whether the MAG or MADG model was adopted, the accuracy of prediction of FER was least, while SL was correspondingly the greatest. The additive variances estimated by the two traits described above are also the smallest and the largest, respectively, at 0.07 and 0.50. A number of studies have indicated that, as the level of heritability increases, the accuracy of genome prediction also increases (39, 56, 57). This was also found in the current study: traits with high heritability, such as SL and FD, showed higher accuracy of prediction than FER and RBC, which, with low heritability, suggested that the level of additive genetic variance has a positive effect on the accuracy of prediction. Moreover, it is seen from **Table 2** that the additive and dominant variances of CFWR were higher than those of FD, but with an accuracy of prediction slightly lower than that of FD (**Table 3**). Interestingly, this was also found

in blood traits; additive and dominant variances of HCT were higher than those of MCH, but the accuracy of prediction of the former was lower. The results above suggested that the dominant variance also has a vital impact on predictability of the model. Alves and coworkers analyzed nine fertility and fecundity traits of dairy cows adopting the model containing additive effects and dominant effects and got similar conclusions: models which include non-additive genetic effects for the majority of traits indicate that the effects of epistasis, dominance, or a combination of the two are as important as the additive effect and occasionally contribute substantially more than the additive effect (8). In this study, for those traits with dominant variance, both additive and dominant variances affect the accuracy of prediction. From the comparison of CFWR and FD, it can be explained that dominant effects may have an impact on the accuracy of prediction greater than that of additive effects in some traits.

For these two models (MAG and MADG), the accuracy of prediction of the MAG model was slightly higher than that of MADG, for both wool and blood traits. The proportion of full-sib relatives and the relationship coefficient of the dominant effect between the training and verification datasets are very small because samples were selected from the same generation with different dams. The MADG model, which combined additive and dominant genetic effects, did not exhibit a clear advantage over the MAG model with only the additive genetic effect, possibly a reflection of only a small proportion of the dominant variance effect information being transferred from the training group to the verification group through 5-fold cross validation. This has also been found in previous studies: Ertl and his colleagues found by genomic analysis of Fleckvieh cattle that when the whole sib relationship between training and validation datasets is small, the accuracy of prediction of the total genetic value by cross validation is not higher in the dominant model than in the additive model (11). In the simulation study of Varona et al. (58), only the population with full-sibs or full-sib offspring could capture the change of breeding value when the additive model was converted to a dominant model (58). The samples collected were selected from the whole Alpine Merino herd in the present study and differed from the Holstein cow dataset that had a large proportion of full-sib relatives. Estimated results from full-sib offspring would not therefore represent the whole population. The results of the present study demonstrated that, compared with the MAG model, the MADG model which included the dominant effect displayed no apparent advantage in terms of predictability, not surprising since the additive genetic effects account for a proportion of the epistasis and dominant effects (8). The inclusion of non-additive genetic effects, therefore, in model fitting is not consistently advantageous (10, 13, 59), which is in agreement with the majority of previous studies (60–62).

For the accuracy of GP, the MADG model containing dominant genetic effects does not show a significant advantage over the MA model with only additive effects. Although the advantage of the MA model over MAD in these traits involved in the current study was not obvious, it could not be ignored. It could be expected that when a larger reference dataset was adopted, especially when it contains more offspring animal records, comparing the predictive ability of models with additive effects and models with additive and dominant effects will get greater benefits for the GP; meanwhile, it is not excluded that genetic models containing additive and dominant effects may be beneficial to the development of specific capabilities of integration for other important traits (34). Furthermore, a number of studies have suggested that there exists a dependent relationship between additive and dominant effects (63, 64). Although the processing of these relationships was somewhat complicated and the calculations substantial, we will collect a larger dataset and attempt to take these factors into consideration in subsequent studies.

CONCLUSIONS

In summary, the present study estimated the additive and dominant variances of 9 traits of Alpine Merino sheep based on two different GBLUP models and used 5-fold cross validation to evaluate the accuracy of prediction of breeding values for these traits. This was the first time GP has been applied to the domesticated Alpine Merino sheep population. Dominant genetic effects account for a large proportion of total phenotypic variation in particular traits (FER, RBC, HCT). Both additive and dominant variances play a vital role in the accuracy of prediction, while for some traits the latter may have a greater impact on the accuracy of prediction than the former. In addition, this study indicates that the predictive capability was not improved when dominant effects were included in the model if the proportion of full-sib relatives in the population was small.

Based on the current results, we will expand the scale of the dataset in the subsequent study and continue to research GP by adding other important traits of the Alpine Merino sheep population, in order to provide more theoretical references for the breeding of this sheep population. Moreover, this study adopts genome-wide SNP information to construct additive and dominant relationship matrices; compared with the relationship matrix based on pedigree, the former is obviously more reliable, which suggests the great contribution of genome-wide SNP information in genome selection. Although the individuals involved in this study only included the Alpine Merino sheep population, the heritability estimation results of SL and FD and the accuracy of GP were very close to those reported in the previous literature. Therefore, it is not excluded to extend these study methods and apply them to other breeds of sheep.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://figshare.com/, https://figshare.com/articles/GS_sample/12497993, https://figshare.com/articles/GSQC_bed/12497963.

ETHICS STATEMENT

AUTHOR CONTRIBUTIONS

SZ: conceptualization, formal analysis, and writing of the original draft. SZ and CY: data curation. BY: funding acquisition. HZ and MH: investigation. YY and TG: methodology. SG and BY: project administration and writing of the review and editing. TW and FL: resources. JL and CY: software. HZ: validation. GQ: visualization. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets.2020.573692/full#supplementary-material

REFERENCES

1. Jacobs A, Elmer KR. Alternative splicing and gene expression play contrasting roles in the parallel phenotypic evolution of a salmonid fish. *bioRxiv [Preprint].* (2020). *bioRxiv:2020.2005.2011.087973.* doi: 10.1101/2020.05.11.087973

2. Goddard M. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica.* (2008) 136:245–57. doi: 10.1007/s10709-008-9308-0

3. Meuwissen THE, Hayes BJB, Goddard MEM. Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* (2001) 157:1819–29.

4. Wiggans GR, Cole J, Hubbard S, Sonstegard T. Genomic selection in dairy cattle: the USDA experience. *Annu Rev Anim Biosci.* (2016) 5:309–27. doi: 10.1146/annurev-animal-021815-111422

5. Miglior F, Fleming A, Malchiodi F, Brito L, Martin P, Baes C. A 100-year review: identification and genetic selection of economically important traits in dairy cattle. *J Dairy Sci.* (2017) 100:10251–71. doi: 10.3168/jds.2017-12968

6. Sun C, Vanraden PM, Cole JB, O'connell JR. Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. *PLoS ONE.* (2014) 9:e103934. doi: 10.1371/journal.pone.0103934

7. Weller J, Ezra E, Ron M. Invited review: a perspective on the future of genomic selection in dairy cattle. *J Dairy Sci.* (2017) 100:8633–44. doi: 10.3168/jds.2017-12879

8. Alves K, Brito L, Baes C, Sargolzaei M, Robinson J, Schenkel F. Estimation of additive and non-additive genetic effects for fertility and reproduction traits in North American holstein cattle using genomic information. *J Anim Breed Genet.* (2020) 137:316–30. doi: 10.1111/jbg.12466

9. Varona L, Legarra A, Toro M, Vitezica Z. Review: non-additive effects in genomic selection. *Front Genet.* (2018) 9:78. doi: 10.3389/fgene.2018.00078

10. Aliloo H, Pryce J, Gonzalez-Recio O, Cocks B, Hayes B. Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. *Genet Select Evol.* (2016) 48:8. doi: 10.1186/s12711-016-0186-0

11. Ertl J, Legarra A, Vitezica Z, Varona L, Edel C, Emmerling R, et al. Genomic analysis of dominance effects on milk production and conformation traits in fleckvieh cattle. *Genet Sel Evol.* (2014) 46:40. doi: 10.1186/1297-9686-46-40

12. Aliloo H, Pryce J, Gonzalez-Recio O, Cocks B, Goddard ME, Hayes BJ. Including nonadditive genetic effects in mating programs to maximize dairy farm profitability. *J Dairy Sci.* (2017) 100:1203–22. doi: 10.3168/jds.2016-11261

13. Munoz Del Valle P, Resende M, Gezan S, Resende M, Campos G, Kirst M, et al. Unraveling additive from non-additive effects using genomic relationship matrices. *Genetics.* (2014) 198:1759–68. doi: 10.1534/genetics.114.171322

14. Gengler N, Van Vleck L, Macneil M, Misztal I, Pariacote F. Influence of dominance relationships on the estimation of dominance variance with sire-dam subclass effects. *J Anim Sci.* (1997) 75:2885–91. doi: 10.2527/1997.75112885x

15. Palucci V, Schaeffer L, Miglior F, Osborne V. Non-additive genetic effects for fertility traits in canadian holstein cattle (open access publication). *Genet Select Evol.* (2007) 39:181–93. doi: 10.1186/1297-9686-39-2-181

16. Norris D, Varona L, Ng'ambi J, Visser D, Mbajiorgu C, Voordewind S. Estimation of the additive and dominance variances in SA duroc pigs. *Livest Sci.* (2010) 131:144–7. doi: 10.1016/j.livsci.2010.03.005

17. Taylor J, Taylor K, Decker J. Holsteins are the genomic selection poster cows. *Proc National Acad Sci USA.* (2016) 113:7690–2. doi: 10.1073/pnas.1608144113

18. Kim J, Lee Y-M, Zahangir Alam M, Dang C-G, Cho K-H, Park K-D. The effectiveness of genomic selection for milk production traits of holstein dairy cattle. *Asian Austr J Anim Sci.* (2019) 33:382–9. doi: 10.5713/ajas.19.0546

19. Taylor J, McKay S, Rolf M, Ramey H, Decker J, Schnabel R. Genomic selection in beef cattle. In: Womack JE, editor. *Bovine Genomics.* (2012). p. 211–33. doi: 10.1002/9781118301739.ch14

20. Samoré A, Fontanesi L. Genomic selection in pigs: state of the art and perspectives. *Ital J Anim Sci.* (2016) 15:211–32. doi: 10.1080/1828051X.2016.1172034

21. Carillier-Jacquin C, Larroque H, Robert-Granié C. Toward genomic selection in dairy goats. *Prod Anim.* (2017) 30:19–30. doi: 10.20870/productions-animales.2017.30.1.2228

22. Zhao Z, Zhang L. Applications of genome selection in sheep breeding. *Hereditas.* (2019) 41:293–303. doi: 10.16288/j.yczz.18-251

23. Toro M, Varona L. A note on mate allocation for dominance handling in genomic selection. *Genet Select Evol.* (2010) 42:33. doi: 10.1186/1297-9686-42-33

24. Su G, Christensen O, Ostersen T, Henryon M, Lund M. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS ONE.* (2012) 7:e45293. doi: 10.1371/journal.pone.0045293

25. Vitezica Z, Varona L, Legarra A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics.* (2013) 195:1223–30. doi: 10.1534/genetics.113.155176

26. Xiaoming M, Jia C, Fu D, Chu M, Ding X, Wu X, et al. Analysis of hematological traits in polled yak by genome-wide association studies using individual SNPs and haplotypes. *Genes.* (2019) 10:463. doi: 10.3390/genes10060463

27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* (2007) 81:559–75. doi: 10.1086/519795

28. Vanraden PM. Efficient methods to compute genomic predictions. *J Dairy Sci.* (2008) 91:4414–23. doi: 10.3168/jds.2007-0980

29. Legarra A, Robert-Granié C, Manfredi E. Performance of genomic selection in mice. *Genetics.* (2008) 180:611–8. doi: 10.1534/genetics.108.088575

30. Daetwyler HD, Swan A, Werf J, Hayes B. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed data assessed by cross-validation. *Genet Select Evol.* (2012) 44:33. doi: 10.1186/1297-9686-44-33

31. Hayes B, Visscher P, Goddard M. Increased accuracy of selection by using the realised relationship matrix. *Genet Res.* (2009) 91:47–60. doi: 10.1017/S0016672308009981

32. Vanraden P, Tassell C, Wiggans GR, Sonstegard T, Schnabel R, Taylor J, et al. Invited review: reliability of genomic predictions for north American holstein bulls. *J Dairy Sci.* (2009) 92:16–24. doi: 10.3168/jds.2008-1514

33. Harris B, Johnson D. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *J Dairy Sci.* (2010) 93:1243–52. doi: 10.3168/jds.2009-2619

34. Su G, Madsen P, Nielsen U, Mäntysaari E, Aamand G, Christensen OF, et al. Genomic prediction for nordic red cattle using one-step and selection index blending. *J Dairy Sci.* (2012) 95:909–17. doi: 10.3168/jds.2011-4804

35. Visscher P, Yang J, Goddard M. A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). *Twin Res Hum Genet.* (2010) 13:517–24. doi: 10.1375/twin.13.6.517

36. Yang J, Benyamin B, Mcevoy B, Gordon S, Henders A, Nyholt D, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* (2010) 42:565–9. doi: 10.1038/ng.608

37. Estellé J, Gil F, Vázquez J, Latorre R, Ramírez Zarzosa G, Barragán M, et al. A QTL genome scan for porcine muscle fiber traits reveals overdominance and epistasis. *J Anim Sci.* (2008) 86:3290–9. doi: 10.2527/jas.2008-1034

38. Duthie C, Simm G, Doeschl-Wilson A, Kalm E, Knap P, Roehe R. Epistatic analysis of carcass characteristics in pigs reveals genomic interactions between quantitative trait loci due to additive and dominance genetic effects. *J Anim Sci.* (2010) 88:2219–34. doi: 10.2527/jas.2009-2266

39. Goddard M, Hayes B. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet.* (2009) 10:381–91. doi: 10.1038/nrg2575

40. Hayes B, Bowman P, Chamberlain A, Verbyla K, Goddard M. Accuracy of genomic breeding values in multi-breed populations. *Genet Select Evol.* (2009) 41:51. doi: 10.1186/1297-9686-41-51

41. Habier D, Tetens J, Seefried F, Lichtner P, Thaller G. The impact of genetic relationship on genomic breeding values in German Holstein cattle. *Genet Select Evol.* (2010) 42:5. doi: 10.1186/1297-9686-42-5

42. Clark S, Daetwyler HD, Werf J. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet Select Evol.* (2012) 44:4. doi: 10.1186/1297-9686-44-4

43. Pszczola MJ, Strabel T, Mulder H, Calus M. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J Dairy Sci.* (2012) 95:389–400. doi: 10.3168/jds.2011-4338

44. Fogarty N. Genetic parameters for live weight, fat and muscle measurements, wool production and reproduction in sheep: a review. *Anim Breed Abstr.* (1995) 63:101–43.

45. Safari E, Fogarty N, Gilmour AR. A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. *Livest Prod Sci.* (2005) 92:271–89. doi: 10.1016/j.livprodsci.2004.09.003

46. Miglior F, Burnside E, Kennedy B. Production traits of holstein cattle: estimation of nonadditive genetic variance components and inbreeding depression. *J Dairy Sci.* (1995) 78:1174–80. doi: 10.3168/jds.S0022-0302(95)76735-2

47. Misztal I, Varona L, Culbertson M, Bertrand J, Mabry J, Lawlor T, et al. Studies on the value of incorporating the effect of dominance in genetic evaluations of dairy cattle, Beef cattle and swine. *Biotechnol Agron Soc Environ.* (1998) 2:227–33.

48. Tassell CP, Misztal I, Varona L. Method R estimates of additive genetic, dominance genetic, and permanent environmental fraction of variance for yield and health traits of holsteins. *J Dairy Sci.* (2000) 83:1873–7. doi: 10.3168/jds.S0022-0302(00)75059-4

49. Kijas J, Lenstra J, Hayes B, Boitard S, Porto-Neto L, Cristobal M, et al. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* (2012) 10:e1001258. doi: 10.1371/journal.pbio.1001258

50. Große-Brinkhaus C, Jonas E, Buschbell H, Phatsara C, Tesfaye D, Heinz J, et al. Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc × Pietrain population. *Genet Select Evol.* (2010) 42:39. doi: 10.1186/1297-9686-42-39

51. Carlborg Ö, Kerje S, Schutz K, Jacobsson L, Jensen P, Andersson L. A global search reveals epistatic interaction between QTL for early growth in the chicken. *Genome Res.* (2003) 13:413–21. doi: 10.1101/gr.528003

52. Carlborg Ö, Hocking P, Burt D, Haley C. Simultaneous mapping of epistatic QTL in chickens reveals clusters of QTL pairs with similar genetic effects on growth. *Genet Res.* (2004) 83:197–209. doi: 10.1017/S0016672304006779

53. Daetwyler HD, Hickey J, Henshall J, Dominik S, Gredler-Grandl B, Werf J, et al. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. *Anim Prod Sci.* (2010) 50:1004. doi: 10.1071/AN10096

54. Moghaddar N, Swan A, Werf J. Genomic prediction of weight and wool traits in a multi-breed sheep population. *Anim Prod Sci.* (2014) 54:544–9. doi: 10.1071/AN13129

55. Hamadani A, Ganai N, Khan NN, Shanaz S, Ahmad T. Estimation of genetic, heritability, and phenotypic trends for weight and wool traits in rambouillet sheep. *Small Ruminant Res.* (2019) 177:133–40. doi: 10.1016/j.smallrumres.2019.06.024

56. Nielsen H, Sonesson A, Yazdi H, Meuwissen T. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. *Aquaculture.* (2009) 289:259–64. doi: 10.1016/j.aquaculture.2009.01.027

57. Xu L, Zezhao W, Zhu B, Liu Y, Li H, Bordbar F, et al. Theoretical evaluation of multi-breed genomic prediction in chinese indigenous cattle. *Animals.* (2019) 9:789. doi: 10.3390/ani9100789

58. Varona L, Misztal I, Bertrand J, Lawlor T. Effect of full sibs on additive breeding values under the dominance model for stature in united states holsteins. *J Dairy Sci.* (1998) 81:1126–35. doi: 10.3168/jds.S0022-0302(98)75675-9

59. Xiang T, Christensen O, Vitezica Z, Legarra A. Genomic evaluation by including dominance effects and inbreeding depression for purebred and crossbred performance with an application in pigs. *Genet Select Evol.* (2016) 48:92. doi: 10.1186/s12711-016-0271-4

60. Esfandyari H, Bijma P, Henryon M, Christensen O, Sørensen C. Genomic prediction of crossbred performance based on purebred Landrace and Yorkshire data using a dominance model. *Genet Select Evol.* (2016) 48:40. doi: 10.1186/s12711-016-0220-2

61. Moghaddar N, Werf J. Genomic estimation of additive and dominance effects and impact of accounting for dominance on accuracy of genomic evaluation in sheep populations. *J Anim Breed Genet.* (2017) 134:453–62. doi: 10.1111/jbg.12287

62. Vitezica Z, Reverter A, Herring W, Legarra A. Dominance and epistatic genetic variances for litter size in pigs using genomic models. *Genet Select Evol.* (2018) 50:71. doi: 10.1186/s12711-018-0437-3

63. Wellmann R, Bennewitz J. The contribution of dominance to the understanding of quantitative genetic variation. *Genet Res.* (2011) 93:1–16. doi: 10.1017/S0016672310000649

64. Wellmann R, Bennewitz J. Bayesian models with dominance effects for genomic evaluation of quantitative traits. *Genet Res.* (2012) 94:21–37. doi: 10.1017/S0016672312000018