



HAL
open science

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

Marc Teissier, Helene Larroque, Christèle Robert-Granié

► To cite this version:

Marc Teissier, Helene Larroque, Christèle Robert-Granié. 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. *Journal of Dairy Science*, 2019, 102 (4), pp.3142-3154. 10.3168/jds.2018-15650 . hal-02625689

HAL Id: hal-02625689

<https://hal.inrae.fr/hal-02625689>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



J. Dairy Sci. 102:1–13
<https://doi.org/10.3168/jds.2018-15650>

© 2019, The Authors. Published by FASS Inc. and Elsevier Inc. on behalf of the American Dairy Science Association®.
 This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

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

M. Teissier,* H. Larroque, and C. Robert-Granie

GenPhySE, Université de Toulouse, INRA, INPT, ENVT, 31326 Castanet-Tolosan, France

ABSTRACT

Genomic evaluation of French dairy goats is routinely conducted using the single-step genomic BLUP (ssGBLUP) method. This method has the advantage of simultaneously using all phenotypes, pedigrees, and genotypes. However, ssGBLUP assumes that all SNP explain the same amount of genetic variance, which is unlikely in the case of traits whose major genes or QTL are segregating. In this study, we investigated the effect of weighted ssGBLUP and its alternatives, which give more weight to SNP associated with the trait, on the accuracy of genomic evaluation of milk production, udder type traits, and somatic cell scores. The data set included 2,955 genotyped animals and 2,543,680 pedigree animals. The number of phenotypes varied with the trait. The accuracy of genomic evaluation was assessed on 205 genotyped Alpine and 146 genotyped Saanen goats born between 2009 and 2012. For traits with unknown QTL, weighted ssGBLUP was less accurate than, or as accurate as, ssGBLUP. For traits with identified QTL (i.e., QTL only present in the Saanen breed), weighted ssGBLUP outperformed ssGBLUP by between 2 and 14%.

Key words: genomic evaluation, quantitative trait loci, weighted single-step genomic best linear unbiased predictor, French dairy goat

INTRODUCTION

Genomic evaluation is routinely used in an increasing number of species including dairy cattle (Hayes et al., 2009; Boichard et al., 2012), poultry (Wolc et al., 2016), dairy sheep (Duchemin et al., 2012; Baloche et al., 2014), meat sheep (Auvray et al., 2014; Brito et al., 2017), pigs (Christensen et al., 2012), and dairy goats (Carillier et al., 2013, 2014; Mucha et al., 2015).

Several genomic methods have been tested and implemented but the most widely used is single-step genomic BLUP (ssGBLUP; Legarra et al., 2009). Single-step genomic BLUP has the advantage of simultaneously using the phenotypes of genotyped and nongenotyped animals, pedigrees, and genotypes. The method constructs a relationship matrix based on the numerator relationship matrix (**A**) and the genomic relationship matrix (**G**) called the hybrid relationship matrix (**H**). The use of ssGBLUP increases the accuracy of genomic evaluation in many contexts and species compared with pedigree-based BLUP or genomic BLUP (**GBLUP**; Chen et al., 2011; Carillier et al., 2014; Onogi et al., 2015; Matilainen et al., 2016). However, the expected increase in the accuracy of genomic evaluation depends on several parameters including the size of the reference population (Lourenco et al., 2014; Andonov et al., 2017), the relationship between the training and validation population (Meuwissen et al., 2001), the extent of linkage disequilibrium (**LD**; Zhou et al., 2018), or the genetic architecture of the trait concerned (Goddard, 2009; Carillier-Jacquin et al., 2016; Zhou et al., 2018).

The main French dairy goats breeds are Alpine and Saanen, and their standard evaluation is based on milk production traits, udder type traits, and SCS. Carillier et al. (2013, 2014) investigated the feasibility of genomic evaluation of French dairy goats using the goat SNP50 BeadChip (Illumina Inc., San Diego, CA). These authors showed that LD is less extensive than in dairy cattle (Carillier et al., 2013), and that the reference population is limited in the Alpine and Saanen breeds. In 2016, the reference population consisted of 2,955 genotyped animals (2,050 females and 905 males; Teissier et al., 2018). Carillier et al. (2013, 2014) concluded that ssGBLUP was more accurate than either the pedigree-based BLUP or GBLUP. Using a multi-breed approach, the authors reported a –4 to 39% change in accuracy for milk production traits using ssGBLUP compared with a pedigree-based BLUP, a 61 to 96% gain in accuracy for udder type traits, and a 54% gain in accuracy for SCS.

Received September 5, 2018.

Accepted December 5, 2018.

*Corresponding author: marc.teissier@inra.fr

In another study, GWAS analyses were performed in dairy goats (French Alpine and Saanen, or mixed-breed goats) to identify QTL that affect traits under selection. Martin et al. (2018) and Mucha et al. (2018) investigated the genetic architecture of different traits. These authors observed a large QTL associated with milk yield, fat yield, protein yield, udder floor position, rear udder attachment, and SCS on chromosome 19. For the standard traits, they also identified important genomic regions on different chromosomes in all breeds or in only one breed. In both French dairy goat breeds, the α_{S1} casein gene associated with protein content was found to be located on chromosome 6 (Grosclaude et al., 1987) and the *DGAT1* gene associated with fat yield on chromosome 14 (Martin et al., 2017). Using these results, our aim was to investigate whether using information on the location of the detected QTL would improve the accuracy of genomic evaluation in an appropriate ssGBLUP. To this end, we tested the incorporation of previous analyses of the effect of the α_{S1} casein gene in the genomic evaluation method (Carillier-Jacquin et al., 2016; Teissier et al., 2018): gene content (Gengler et al., 2007; Legarra and Vitezica, 2015), weighted ssGBLUP (**WssGBLUP**; Wang et al., 2012), WssGBLUP alternatives (Zhang et al., 2016), and TABLUP (Zhang et al., 2015). Gene content is a multiple-trait ssGBLUP model in which the genotype for a specific causal mutation is considered as a new trait, thus enabling the combination of information from SNP and genotypes for a causal mutation. It can be extended to multi-allelic genes and used when a causal mutation is missing. The WssGBLUP and alternatives are based on the ssGBLUP framework in which weights for SNP variances are used to form the genomic relationship matrix **G**. The WssGBLUP can give more weight to SNP that are in high LD with a causal mutation or associated with QTL with a relatively large effect. The weights were estimated from the variance explained by each SNP as described by Wang et al. (2012). **G** is trait specific and depends on the genetic architecture of the trait (traits with QTL or polygenic traits). With WssGBLUP, one weight is allocated to each SNP, whereas alternative WssGBLUP use the same weight for SNP that are located within a defined window along the genome (Zhang et al., 2016). With alternatives WssGBLUP, the weight in a defined window is calculated as the sum of all SNP weights of the window (**WssGBLUP_{Sum}**) or as the maximum of the SNP weights of the window (**WssGBLUP_{Max}**). Finally, TABLUP is ssGBLUP with a genomic relationship matrix based on genotypes from a subset of pre-selected SNP. The SNP can be selected after GWAS analysis or based on

weights estimated with WssGBLUP. The selected SNP are then given equal weights for the analyses (Zhang et al., 2011). Carillier et al. (2016) and Teissier et al. (2018) showed that only WssGBLUP and their alternatives are able to outperform pedigree-based BLUP and ssGBLUP (Teissier et al., 2018). Compared with ssGBLUP, neither the gene content method nor TABLUP on protein content improved the accuracy of genomic evaluations in either breed. On the other hand, with WssGBLUP and their alternatives (WssGBLUP_{Sum} or WssGBLUP_{Max}), improvements were observed, with +6 percentage points of accuracy over ssGBLUP. The advantage of WssGBLUP and alternative WssGBLUP over the gene content method is that only genotypes from SNP50 BeadChip are required.

The aim of this study was to investigate the use of WssGBLUP and WssGBLUP alternatives for other widely selected traits with different genetic architectures and, in some cases, with QTL identified as having a relatively large effect. The accuracy of WssGBLUP methods and ssGBLUP were compared. The weights of SNP and their effect on the genomic relationship matrix were investigated for all the traits. The effect of this weighting on the accuracy of genomic evaluation was also investigated and compared with that obtained with the ssGBLUP method.

MATERIALS AND METHODS

Pedigree, Genotyped, and Phenotyped Animals

The data sets included phenotypes, pedigree, genotypes (Illumina goat SNP50 BeadChip), and environmental effects of the 2 main French dairy goat breeds (Alpine and Saanen) obtained from the French National Milk Recording System (<http://fr.france-genetique-elevage.org>). Data from the official genetic evaluation in January 2016 were used in this study. All analyses were within breed.

The standard traits selected in French dairy goats include 4 milk production traits, milk yield (**MY** in kg), fat and protein yields (**FY** and **PY**, respectively, in kg), and fat content (**FC** in g/kg), 5 udder type traits, teat angle (**TA**: scored from 1 to 9), udder floor position (**UFP**: scored from 1 to 9), rear udder attachment (**RUA**: scored from 1 to 9), fore udder (**FU**: scored from 1 to 9), and udder shape (**US**: scored from 1 to 9), and SCS (log-transformed SCC) were analyzed. Descriptive statistics on the number of records, and the mean and the heritability of each trait and each breed studied are presented in Table 1. Milk production traits were expressed as 250-d yields. Almost 4 million phe-

Table 1. Number of records, mean, and heritability (h^2) for the traits studied in the Alpine and Saanen breeds

Item	Alpine			Saanen		
	Performances (no.)	Mean	h^2	Performances (no.)	Mean	h^2
Milk yield ¹ (kg)	3,844,314	802.12	0.31	2,923,531	823.08	0.26
Fat yield ¹ (kg)	3,742,129	28.4	0.28	2,887,051	27.44	0.25
Protein yield ¹ (kg)	3,844,071	24.36	0.31	2,923,419	24.32	0.25
Fat content ¹ (g/kg)	3,742,129	35.33	0.48	2,887,051	33.39	0.51
Teat angle (score)	150,676	3.63	0.42	102,967	4.05	0.45
Udder floor position (score)	150,676	6.37	0.51	102,967	6.16	0.57
Rear udder attachment (score)	150,676	4.57	0.47	102,967	4.96	0.52
Fore udder (score)	150,676	3.19	0.44	102,967	3.38	0.42
Udder shape (score)	150,676	5.76	0.40	102,967	6.22	0.47
SCS	1,262,187	165.03	0.20	1,031,450	158.99	0.16

¹Expressed as 250-d yields.

notypes in Alpine and 3 million phenotypes in Saanen were recorded for milk production traits. More than 150,000 phenotypes were available in the Alpine breed and more than 100,000 phenotypes in the Saanen breed for udder type traits, and 1.2 million phenotypes in the Alpine breed and 1 million phenotypes in the Saanen breed for SCS.

The pedigree file contained animals born between 1936 and 2012. For milk production traits, 1,446,296 Alpine animals and 1,097,384 Saanen animals were used. For udder type traits, the pedigree file included 290,656 Alpine animals and 206,154 Saanen animals. For SCS, the pedigree file contained 788,576 Alpine animals and 648,461 Saanen animals. The pedigree file was then completed with unknown parent groups: one group was created for animals born before 1975 and then pooled groups (sires and dams) were defined every 2 yr. Males and females were pooled together in unknown parent groups because few animals had unknown dams.

French dairy goats were genotyped with the Illumina goat SNP50 BeadChip (50K SNP; Tosser-Klopp et al., 2014). Quality control (QC) was applied to 2,056 genotyped Alpine and 1,349 genotyped Saanen animals (born between 1983 and 2012) for 53,347 SNP, independently for each breed. During the QC, SNP with a minor allele frequency (i.e., less than 1% and a call rate of less than 95%) were removed. The Hardy-Weinberg equilibrium for each SNP was tested by calculating the associated chi-squared statistic. The SNP with a P -value lower than 1.10^{-6} were removed (threshold of 5% corrected for multiple testing). Animals with a SNP call rate of less than 90% were discarded from the analyses. Finally, parent-progeny Mendelian conflicts were checked. After the QC, 1,749 genotyped Alpine (512 males and 1,237 females) and 1,206 genotyped Saanen (393 males and 813 females) for 46,849 SNP remained for analyses. These animals were born between 1993 and 2012.

ssGBLUP

The ssGBLUP is a routinely used method for genomic evaluation of 11 traits selected in the 2 main French dairy goats (Carillier et al., 2014; Venot et al., 2017). It can simultaneously combine information on female phenotypes, pedigrees, and genotypes. Each trait was analyzed with a single trait model. For milk production traits (MY, FY, PY, FC) and SCS, the same model as in the routine genetic evaluation was used (Clément et al., 2002):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{p} + \mathbf{e}, \quad [\text{model 1}]$$

where \mathbf{y} is a vector of phenotypes, $\boldsymbol{\beta}$ is a vector of fixed effects including 4 combined effects: herd, age and month at kidding, and length of the dry period. The herd effect was estimated within year (32 yr from 1980 to 2012) and parity (1, 2, and ≥ 3); age and month were within year and region estimations (4 regions in France depending on goat breeding management). The length of the dry period was an estimation within a year and region. \mathbf{u} is a vector of genomic breeding values (GEBV) assumed to be normally distributed $N(\mathbf{0}, \mathbf{H}\sigma_u^2)$, where \mathbf{H} represents the relationship matrix and σ_u is the variance of the random additive genetic effect, \mathbf{p} is a vector of random permanent environmental effects assumed to be normally distributed $N(\mathbf{0}, \mathbf{I}\sigma_p^2)$, with σ_p the variance of the permanent environmental effect, and \mathbf{e} is a vector of random residual normally distributed $N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, with σ_e the variance of residuals. \mathbf{X} is the incidence matrix relating phenotypes to fixed effects ($\boldsymbol{\beta}$); \mathbf{Z} is the design matrix which allocates phenotypes to genomic breeding values (\mathbf{u}) and \mathbf{W} is the incidence matrix that links phenotypes to permanent environmental effects (\mathbf{p}). Solutions of $\boldsymbol{\beta}$, \mathbf{u} , and \mathbf{p} were obtained by solving the following system:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \frac{\sigma_e^2}{\sigma_u^2}\mathbf{H}^{-1} & \mathbf{Z}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{Z} & \mathbf{W}'\mathbf{W} + \frac{\sigma_e^2}{\sigma_p^2}\mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}.$$

A different model was used to analyze udder type traits. The only difference is that no permanent environmental effect was estimated because the animals were scored only once in their life (during their first parity). The model was the following:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \quad [\text{model 2}]$$

where \mathbf{y} , \mathbf{u} , and \mathbf{e} are the same vectors previously described in model 1 and $\boldsymbol{\beta}$ is the vector of 3 combined fixed effects: herd, age at scoring, and stage at scoring. Herd effect and parity, age at scoring, and stage at scoring were within year estimations. The matrix \mathbf{H}^{-1} is expressed as

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

which depend on the inverse of the relationship matrix estimated from the pedigree \mathbf{A} (subscript 22 refers to genotyped animals) and on the inverse of the genomic relationship matrix \mathbf{G} . The \mathbf{G} matrix was estimated using genotypes as in Legarra et al. (2009) or Misztal et al. (2013):

$$\mathbf{G} = 0.95 \frac{\mathbf{M}'\mathbf{M}}{2\sum_{i=1}^m p_i(1-p_i)} + 0.05\mathbf{A}_{22},$$

where m is the number of SNP, p_i is the estimated allele frequency at the locus i , and \mathbf{M} is a centered matrix of SNP genotypes.

Variance components were estimated using the REML method in the remlf90 software and ssGBLUP analyses were performed with the blup90iod2 software (Misztal et al., 2002).

Weighted ssGBLUP

The construction of the \mathbf{G} matrix presented above assumes that each SNP explains the same amount of genetic variance. Consequently, this assumption is not

valid for traits with a major gene or QTL. Wang et al. (2012) proposed another genomic approach called WssGBLUP based on a model similar to ssGBLUP, to include major genes or QTL with a relatively large effect using a weighted \mathbf{G} (\mathbf{G}^*). This genomic relationship matrix \mathbf{G}^* is constructed as follows:

$$\mathbf{G}^* = 0.95 \frac{\mathbf{M}'\mathbf{D}\mathbf{M}}{2\sum_{i=1}^m p_i(1-p_i)} + 0.05\mathbf{A}_{22},$$

where \mathbf{A}_{22} , \mathbf{M} , p_i , and m are the same as in \mathbf{G} and \mathbf{D} is a diagonal matrix of size $m \times m$, where each element of the diagonal corresponds to SNP weights.

The WssGBLUP approach is based on an iterative algorithm with different steps: (1) run ssGBLUP with the \mathbf{G}^* matrix (at iteration 1, SNP weights in the \mathbf{D} matrix are equal to 1 and are equivalent to a ssGBLUP), (2) estimate SNP effects from solutions of genomic breeding values in the previous step, (3) estimate variances of the effect of each SNP, (4) normalize the vector of variances of SNP effects to get the SNP weights (this normalization process ensures that the sum of the variances remain constant and equal to the number of SNP), (5) use SNP weights to construct the \mathbf{D} matrix, and (6) loop to step (1).

The WssGBLUP was applied to each trait studied with model 1 and 2, respectively, using blupf90 family software (blup90iod2, Misztal et al., 2002). The SNP effects and SNP weights were estimated using postGSf90 software. In this study, 3 WssGBLUP approaches were investigated, each one using a specific \mathbf{G}^* : WssGBLUP, WssGBLUP_{Sum}, and WssGBLUP_{Max}. The WssGBLUP is the method presented by Wang et al. (2012), which consists in attributing one weight to each SNP. With WssGBLUP_{Sum} and WssGBLUP_{Max} (Zhang et al., 2016), SNP on the whole genome are split into different-sized nonoverlapping windows, and the same weight is given to each SNP of the window. Windows of 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 consecutive SNP were tested. To compute these weights, the sum of all SNP weights present in the window was given to those SNP (WssGBLUP_{Sum}), or the SNP with the highest weight in the window was given to all SNP in the same window (WssGBLUP_{Max}). The final step consists in normalizing the new vector of SNP weights to ensure that the sum of the variances remains constant and equal to the number of SNP. A previous study has shown that the 2nd iteration of the WssGBLUP with 40 SNP was the most accurate for WssGBLUP_{Max} and WssGBLUP_{Sum} (Teissier et al., 2018), and results were presented for this scenario.

Accuracy of the Genomic Evaluation

The ssGBLUP used female phenotypes, pedigrees, and genotypes. In the French Alpine and Saanen dairy goat breeding scheme, genetic selection is performed on progeny-tested bucks and all these bucks born after 1993 were genotyped. The reference population used to assess the accuracy of genomic evaluation comprised only genotyped males even if genotypes of females were also used in ssGBLUP and WssGBLUP evaluations. This reference population was split into 2 subsets: a training set and a validation set. The training population included 307 Alpine bucks and 247 Saanen bucks born between 1993 and 2007, all the information on these animals (genotype, the pedigree of their ancestry and their progeny, and the phenotypes of their progeny) was kept in the data sets to estimate GEBV. The validation set included 205 Alpine bucks and 146 Saanen bucks born between 2008 and 2012. For these animals, the phenotypes of their progeny were removed from the analysis, and only the genotypes and pedigree of their ancestry were retained. The accuracy of genomic evaluation was calculated as the Pearson correlation between GEBV estimated using the validation set and daughter yield deviations (DYD) calculated using the official genetic evaluation of January 2016. The number of daughters used to calculate these DYD was between 46 and 2,509 (with a median of 177 daughters), indicating that the DYD were relatively accurate. Accuracies of genomic evaluations were compared between ssGBLUP and WssGBLUP and its alternatives with the Hotelling-Williams test (Van Sickle, 2003).

Relationship Coefficients Estimated Using Pedigree and Genomic Information

Elements of the off-diagonal of the numerator relationship matrix for genotyped animals (\mathbf{A}_{22}) and the weighted genomic relationship matrix (\mathbf{G}^*) were compared. To this end, the Pearson correlation between the 2 vectors was calculated.

RESULTS AND DISCUSSION

Accuracy with WssGBLUP Over Iterations

WssGBLUP is based on an iterative process. The first iteration corresponds to ssGBLUP; SNP weights are all equal to 1. The accuracy of the genomic evaluations from iteration 2 to iteration 4 for each breed and each trait are presented in Figure 1. The average accuracies for the 10 traits in the Alpine and Saanen breeds were, respectively, 0.42 and 0.51 at iteration 2, 0.39 and 0.49 at iteration 3, and 0.35 and 0.44 at iteration 4. For all

traits in the Alpine breed, accuracy at iteration 2 was higher than the accuracy at iteration 3, which in turn was higher than accuracy at iteration 4. In the Saanen breed, MY accuracy increased at iteration 3 (0.58) compared with iteration 2 (0.56) and then decreased at iteration 4 (0.52). For (FY), accuracy at iteration 2 and 3 was 0.47, then decreased to 0.43 at iteration 4. For all the other traits (i.e., PY, FC, TA, UFP, RUA, FU, US, and SCS), accuracy decreased over the 3 iterations. In both breeds, the decrease in accuracy between iterations 3 and 4 was bigger than the decrease in accuracy between iterations 2 and 3.

In a previous study, we investigated a similar approach (WssGBLUP) to the analysis of protein content in the same 2 French dairy goat populations (Teissier et al., 2018). We concluded that WssGBLUP at iteration 2 provided the most accurate genomic evaluation. In the present study, we obtained the same results for all the standard traits selected in French national genomic evaluations. Our results are also in agreement with those of Wang et al. (2012), who reported WssGBLUP produced the most accurate genomic evaluation at iteration 2. However, after iteration 2, loss of accuracy in our study was much greater than that observed by Wang et al. (2012). These differences could be due to the fact that Wang et al. (2012) used simulation in their study to mimic a trait with a phenotypic mean of 5, variance of 1, and heritability of 0.5. They simulated 2 chromosomes each with 15 QTL sampled from a gamma distribution with a shape factor of 0.4 and a scale factor of 1. They repeated the simulation 10 times. Overall, the average effect of the QTL was 0.16 (0.04). In our study, the situation was probably more complex because we analyzed real data concerning traits with different genetic architectures. According to Wang et al. (2012) and to our previous study (Teissier et al., 2018), the decrease in accuracy could be due to over/underweighting of some SNP across iterations. In the present study, we consequently investigated this point to check if our SNP weights across the WssGBLUP iterations were inflated.

Effects of Iterations on the WssGBLUP Method

Figure 2 presents the SNP weights for RUA in the Saanen breed between iteration 2 and iteration 4 using WssGBLUP. At iteration 2, the highest SNP weights located on chromosome 19 reached 46. On the whole genome, 95% of SNP had weights under 4. At the 4th iteration with WssGBLUP, SNP weights reached 8,621 on chromosome 19 and 95% of the SNP had weights under 0.14. The results at iteration 3 are not shown but were intermediate between those at iteration 2 and it-

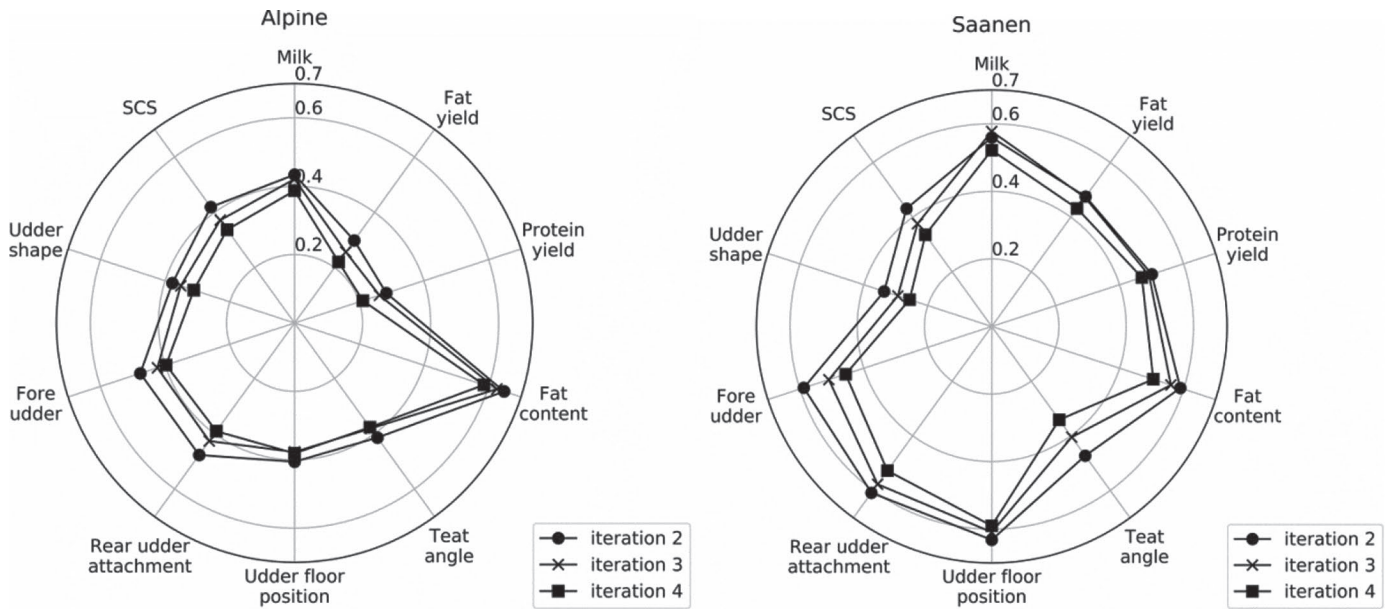


Figure 1. Validation correlations for 205 and 146 validation Alpine and Saanen males, respectively, for 4 milk production traits, 5 udder type traits, and SCS using the weighted single-step genomic BLUP (WssGBLUP) approach at iterations 2, 3, and 4.

eration 4 with a maximum of 1,409 for SNP weights on chromosome 19 with 95% of SNP weights under 2.93. We also observed this huge inflation of SNP weights in the Saanen breed on chromosome 19 for MY, PY, UFP, and SCS. For the other traits (FY, FC, TA, FU, and US), high SNP weights were attributed to some SNP at iteration 4 in both breeds. However, these high SNP weights were not located on a specific chromosome and their maxima were much lower than those observed for MY, PY, UFP, and SCS on chromosome 19 in the

Saanen breed. For instance, in the Alpine breed, the highest SNP weights for RUA reached 2,000 at iteration 4 with WssGBLUP whereas they reached 8,621 in the Saanen breed.

The SNP weights were very highly inflated between 2 iterations; at iteration 4 for RUA in the Saanen breed, 18% of SNP weights were allocated to only one SNP on chromosome 19. For all traits and the 2 breeds, we observed that some SNP strongly associated with the traits considered had very high weights and that SNP

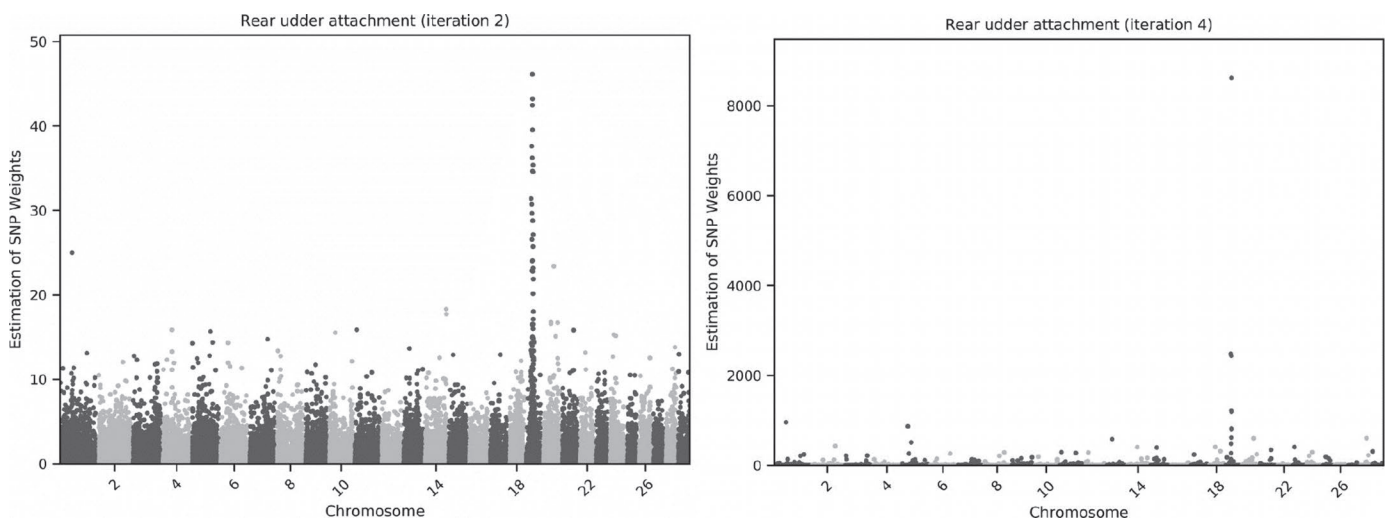


Figure 2. Estimated weights of SNP for rear udder attachment in the Saanen breed at iterations 2 and 4 with the weighted single-step genomic BLUP (WssGBLUP) approach.

weights increased markedly from one iteration to the next, whereas the other SNP weights decreased toward zero. The SNP weights estimated with WssGBLUP were used as weights in matrix \mathbf{D} to construct the weighted matrix \mathbf{G}^* . This matrix, which is included in the \mathbf{H} matrix, could affect the structure of the relationship matrix. We compared off-diagonal elements between \mathbf{G}^* and \mathbf{A}_{22} to observe how much effect SNP weights have on elements of the genomic relationship matrix. Figure 3 shows the correlation between the off-diagonal elements of the \mathbf{G}^* matrix and those of the \mathbf{A}_{22} matrix for each breed and trait. In the Alpine breed, the average correlation for the 10 traits was 0.87 at iteration 2. Few variations were observed among the traits with correlations ranging from 0.84 to 0.91. The best correlations were obtained for milk production traits (MY, FY, PY, and FC). At iteration 3, the average correlation was lower (0.76), range: 0.72 to 0.80. Finally, at iteration 4, we observed a low average correlation (0.47), range: 0.41 to 0.51. In the Saanen breed, similar conclusions were drawn. The average correlation for the 10 traits was 0.80, 0.45, and 0.28 at iterations 2, 3, and 4, respectively, with values ranging from 0.70 to 0.90 at iteration 2, from 0.17 to 0.74 at iteration 3, and from 0.12 to 0.48 at iteration 4.

The profiles of the average correlations in the Alpine and Saanen breeds were similar at iteration 2 but differed markedly at iterations 3 and 4. The loss of

correlation between iteration 2 and 3 in the Alpine breed was roughly equal to 10 percentage points and 30 percentage points between iteration 3 and 4 for any trait included in this study. In the Saanen breed, the loss of correlation was trait dependent and reached 62 percentage points for MY and 7.5 percentage points for US between iteration 2 and 3.

From iteration 3, the off-diagonal elements of \mathbf{G}^* and \mathbf{A}_{22} differ significantly in the Saanen breed. Martin et al. (2018) found QTL for MY, FY, PY, UFP, RUA, and SCS in the French Saanen breed. These traits are those for which we observed the biggest decrease in the correlation between elements of \mathbf{A}_{22} and \mathbf{G}^* . In the Alpine breed, no QTL were detected, suggesting that these traits have polygenic architecture (Martin et al., 2018). We conclude that for those traits for which QTL have been detected, the weights assigned to the SNP most strongly associated with the trait are exacerbated from one iteration to another in the iterative process of WssGBLUP. These results suggest the iterative process of WssGBLUP should be stopped at iteration 2.

Estimation of Weights with WssGBLUP

We analyzed the estimation of SNP weights with WssGBLUP for the 10 traits in both the Alpine and Saanen breeds. These analyses were performed to highlight important chromosomal regions associated with

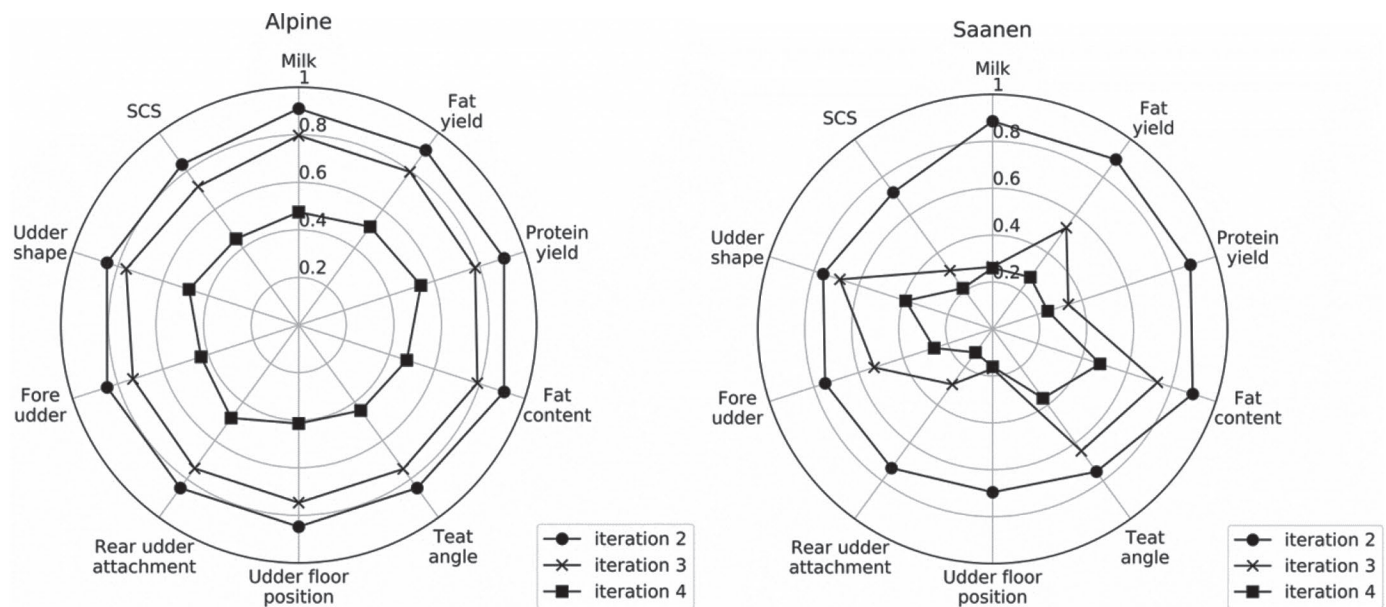


Figure 3. Correlation between off-diagonal elements of the genomic relationship matrix (\mathbf{G}^*) and off-diagonal elements of \mathbf{A}_{22} (pedigree \mathbf{A} with subscript 22 referring to genotyped animals) on 205 and 146 validation Alpine and Saanen males, respectively, with the different iterations of the weighted single-step genomic BLUP (WssGBLUP) for 4 milk production traits, 5 udder type traits, and SCS in Alpine and Saanen populations.

selected traits in French dairy goats. We identified 2 different groups: (1) traits with high SNP weights detected in the Saanen breed on one chromosome but not in the Alpine breed, and (2) traits with SNP weights homogeneously distributed along the chromosomes in the 2 breeds. Figure 4 illustrates these 2 different groups of SNP weights for UFP and US. For UFP (included in the first group), SNP weights were below 30 in the Alpine breed, whereas in the Saanen breed, SNP weights reached 68 for some SNP on chromosome 19. The SNP on other chromosomes had SNP weights below 30. The top 10 SNP with the highest SNP weights on chromosome 19 were located between 26 and 28 Mb. The MY, PY, RUA, and SCS were included in the first group. Except for chromosome 19, the SNP weights for all chromosomes were below 30. On chromosome 19, the maximum weights observed were 48 for MY (top 10 SNP were located between 26 and 29 Mb), 42 for PY (top 10 SNP between 26 and 29 Mb), 46 for RUA (top 10 SNP between 20 and 28 Mb), and 37 for SCS (top 10 SNP between 23 and 28 Mb). The top 10 SNP covered a chromosomal region between 20 and 29 Mb for all these traits. For US (included in group 2), we observed SNP weights below 30 for all SNP in both Alpine and Saanen breeds. The same profile was observed for FY, FC, TA, FU, and US. Martin et al. (2018) performed LD and linkage analysis in French dairy goats. They showed that chromosome 19 underlies a pleiotropic QTL located between 24.5 and 26.9 Mb (5% CI) affecting MY, FY, PY, UFP, and RUA. In our study with WssGBLUP, the highest SNP weights were located on the same chromosome 19 and in the same region but with a slightly larger interval for MY, PY, UFP, and RUA. For FY on chromosome 19, we did not find any SNP with significant weights like those found by Martin et al. (2017). It is possible that our training population was too small, and that with more genotyped animals, we would have identified SNP with higher weights on chromosome 19. Surprisingly, the chromosomal region of DGAT1 on chromosome 14, which is known to be associated with FC, was not identified with WssGBLUP, whereas it was detected by Martin et al. (2017) with LD and linkage analysis. This result shows that SNP effects estimation with WssGBLUP had some limitations and could be improved in the future. This limitation may be due to the whole-genome regression performed to estimate SNP effects, resulting in unstable prediction of SNP effects because of LD between SNP. Martin et al. (2018) found a QTL for SCS in the Saanen breed located between 33 and 42 Mb on chromosome 19. In our study, the highest SNP weights were located on the same chromosome 19 but in a neighboring chromosomal region (between 23 and 28 Mb). In a previous

study Teissier et al. (2018), we conducted WssGBLUP analysis for protein content in which a major gene (α_{S1} casein gene) was identified, but no SNP in the α_{S1} casein gene was on the 50K SNP after QC. We identified some SNP with high weights (between 90 to 101) in the α_{S1} casein gene region on chromosome 6, and this method provides a more accurate genomic evaluation than ssGBLUP in the 2 breeds. This shows that WssGBLUP is able to capture the complexity of this gene.

Accuracy of Genomic Evaluation Using WssGBLUP

Figure 5 presents the accuracy of genomic evaluation using the validation set for the 10 traits. We compared the ssGBLUP method used as a reference method and the WssGBLUP at iteration 2. In the Alpine breed, accuracy was on average slightly lower with WssGBLUP (0.42) than with ssGBLUP (0.44). The loss of accuracy ranged between +0 percentage points (TA or US) to -3 percentage points (SCS); however, these differences were not significant. For this breed, no QTL was identified for these traits (Martin et al., 2018) and no large SNP weight was identified with WssGBLUP. In the Saanen breed, the accuracy of the genomic evaluation was on average slightly higher with WssGBLUP (0.52) than with ssGBLUP (0.51). However, we observed an increase or a decrease in accuracy depending on the trait. With WssGBLUP, accuracy was the same or lower than with ssGBLUP for FC (+0 percentage points), TA (-1 percentage point), US (-2 percentage points); these differences were not significant. However, for FU (-3 percentage points) and SCS (-3 percentage points) significant decrease of accuracies at 0.05 threshold were observed. Among these traits, all SNP weights were low, below 30, except for SCS where a QTL was identified on chromosome 19 (Martin et al., 2018) and SNP weights were higher on a chromosomal region of chromosome 19. Increased accuracy was obtained for the remaining traits with WssGBLUP: +1 percentage point for RUA (not significant at the 0.05 threshold), +4 percentage points for FY and UFP ($P < 0.05$), +5 percentage points for PY ($P < 0.05$) and +7 percentage points for MY ($P < 0.001$). For these traits, QTL were identified on chromosome 19 (Martin et al., 2017, 2018), and except for FY, high SNP weights were also identified in the same chromosomal region with WssGBLUP.

With WssGBLUP, the accuracy of genomic evaluation was improved for traits with segregation of QTL in French dairy goats. In our study, the Saanen breed was mostly concerned with the large QTL on chromosome 19 for MY, PY, UFP, and RUA. These results are consistent with those in our previous study (Teissier et al.,

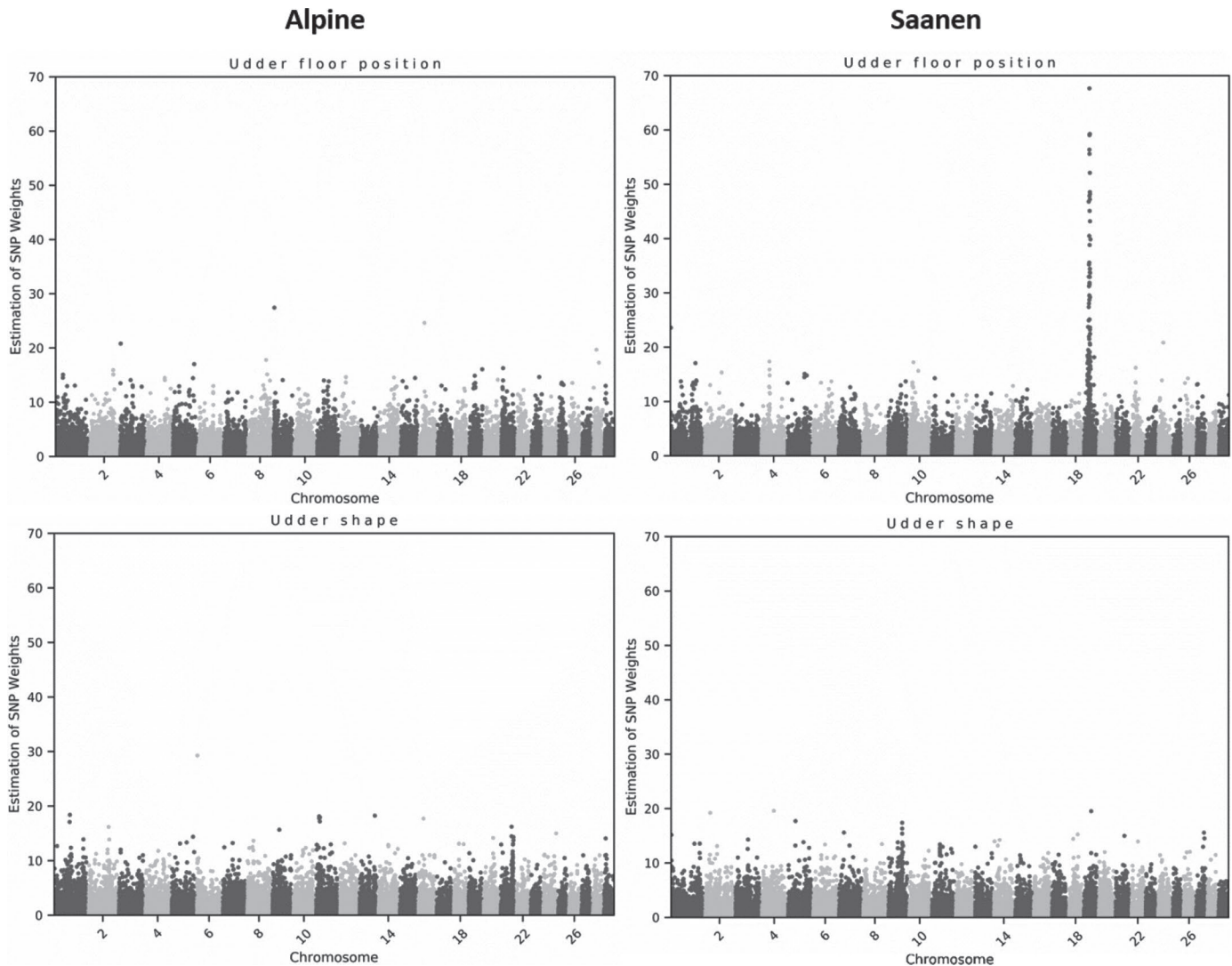


Figure 4. Estimated weights of SNP for udder floor position and udder shape at iteration 2 with weighted single-step genomic BLUP (WssGBLUP) for 4 milk production traits, 5 udder type traits, and SCS in Alpine and Saanen populations.

2018) on protein content, as the α_{S1} casein gene is well known to segregate in these populations. The results of that study showed that with WssGBLUP, accuracy was better than with ssGBLUP in both breeds (+2 and +4 percentage points in the Alpine and Saanen breeds, respectively). Only in Saanen was the WssGBLUP significantly more accurate than ssGBLUP. In another study on dairy cattle using WssGBLUP with a relatively small reference population (1,500 genotyped animals), Lourenco et al. (2014) showed that WssGBLUP could outperform GBLUP or BayesC for traits with QTL with large or moderate effects. On the other side, traits with a polygenic determinism did not benefit from the use of the WssGBLUP method, as a slightly decrease in accuracy was observed compared with ssGBLUP.

The average higher accuracy in Saanen than Alpine breed may be explained by the structure of the population. The level of inbreeding in Saanen (2.3%) is higher than in Alpine (1.8%). There is also a higher kinship coefficient between the training and validation population in the Saanen breed (2.4%) than in Alpine breed (1.1%; Carillier et al., 2013).

Fine Tuning of Weights in the WssGBLUP Method

We observed that with WssGBLUP, SNP weights increased considerably from iteration 2 to iteration 4. Zhang et al. (2016) reported that WssGBLUP alternatives (WssGBLUP_{Sum} and WssGBLUP_{Max}) increased the accuracy of genomic evaluation more efficiently

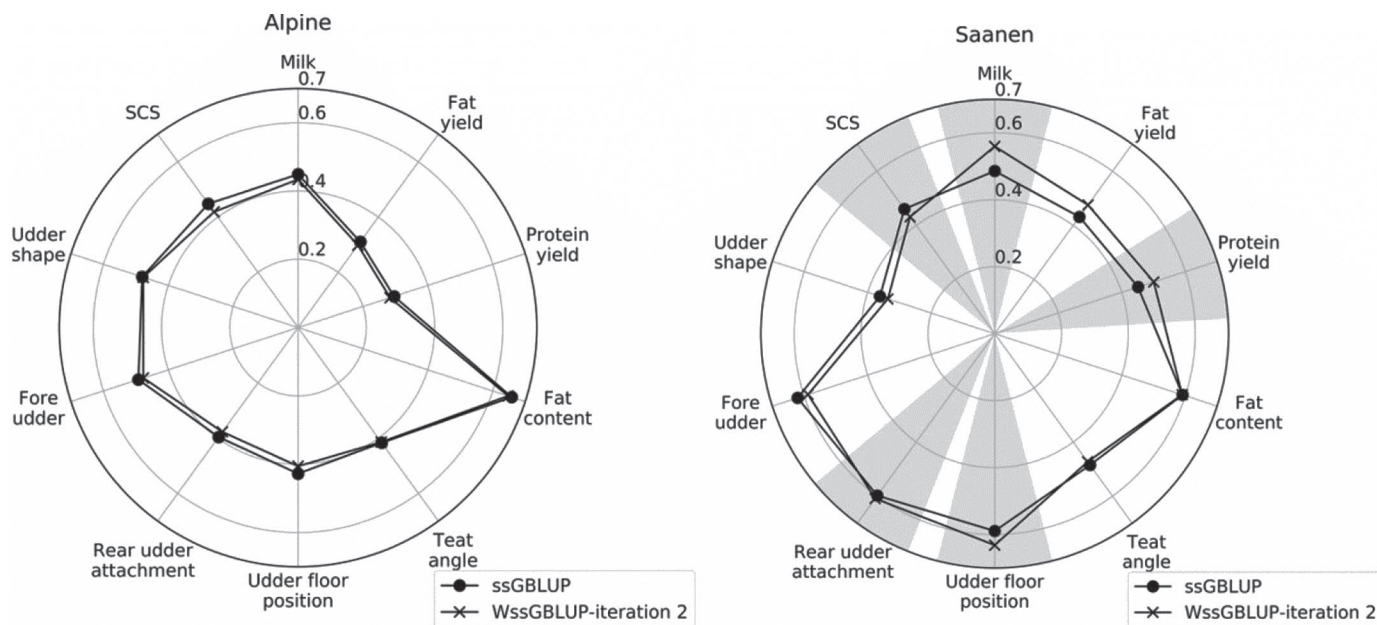


Figure 5. Validation correlations for, respectively, 205 and 146 validation Alpine and Saanen males for 4 milk production traits, 5 udder type traits, and SCS using the single-step genomic BLUP (ssGBLUP) and weighted single-step genomic BLUP (WssGBLUP) at iteration 2. Gray area: traits with high SNP weights detected using WssGBLUP.

than WssGBLUP and limited the increase in SNP weights from one iteration to another. In our previous study (Teissier et al., 2018), we applied $WssGBLUP_{Sum}$ and $WssGBLUP_{Max}$ to protein content and showed that the optimal length of the window was 40 SNP. Even though $WssGBLUP_{Sum}$ and $WssGBLUP_{Max}$ limited the increase in SNP weights over iterations, the best accuracies were obtained at iteration 2. In the present study, we obtained the same results. Table 2 compares the results obtained with ssGBLUP, WssGBLUP, $WssGBLUP_{Sum}$, and $WssGBLUP_{Max}$ with a window size of 40 SNP for the 10 traits at iteration 2 for the Alpine

breed and Table 3 compares the same results for the Saanen breed.

We first compared $WssGBLUP_{Sum}$ with $WssGBLUP_{Max}$. In the Alpine breed, for MY, FY, FC, TA, and UFP, $WssGBLUP_{Sum}$ was as accurate (no significant difference at 0.05 threshold was observed) as $WssGBLUP_{Max}$. For FU and RUA, $WssGBLUP_{Sum}$ was slightly less accurate (−1 percentage point) than $WssGBLUP_{Max}$. For PY, US, and SCS, $WssGBLUP_{Sum}$ was slightly more accurate (+1 percentage point) than $WssGBLUP_{Max}$. In the Saanen breed, for MY, FY, PY, TA, UFP, and SCS, $WssGBLUP_{Sum}$ was as accurate

Table 2. Pearson correlation between genomic breeding values and daughter yield deviations for the traits studied in the Alpine breed¹

Trait	ssGBLUP	WssGBLUP	$WssGBLUP_{Max}$	$WssGBLUP_{Sum}$
Milk yield (kg)	0.45	0.43	0.44	0.44
Fat yield (kg)	0.31	0.30	0.30	0.30
Protein yield (kg)	0.30	0.28	0.28	0.29
Fat content (g/kg)	0.66	0.65	0.66	0.66
Teat angle (score)	0.42	0.41	0.41	0.41
Udder floor position (score)	0.43	0.41	0.44	0.44
Rear udder attachment (score)	0.40	0.38	0.42	0.41
Fore udder (score)	0.49	0.48	0.50	0.49
Udder shape (score)	0.48	0.48	0.48	0.49
SCS	0.45	0.42	0.44	0.45
Mean	0.44	0.42	0.44	0.44

¹ssGBLUP = single-step genomic BLUP; WssGBLUP = weighted single-step genomic BLUP. Accuracies for the maximum of the SNP weights of the window ($WssGBLUP_{Max}$) and the sum of all SNP weights of the window ($WssGBLUP_{Sum}$) are presented for a window size of 40 consecutive SNP.

Table 3. Pearson correlation between genomic breeding values and daughter yield deviations for the traits studied in the Saanen breed¹

Trait	ssGBLUP	WssGBLUP	WssGBLUP _{Max}	WssGBLUP _{Sum}
Milk yield (kg)	0.49	0.56	0.56	0.56
Fat yield (kg)	0.43	0.48	0.49	0.49
Protein yield (kg)	0.45	0.50	0.51	0.51
Fat content (g/kg)	0.59	0.59	0.60	0.61
Teat angle (score)	0.49	0.47	0.48	0.48
Udder floor position (score)	0.59	0.63	0.65	0.65
Rear udder attachment (score)	0.60	0.61	0.62	0.63
Fore udder (score)	0.62	0.59	0.62	0.63
Udder shape (score)	0.36	0.34	0.36	0.37
SCS	0.46	0.43	0.46	0.46
Mean	0.51	0.52	0.54	0.54

¹ssGBLUP = single-step genomic BLUP; WssGBLUP = weighted single-step genomic BLUP. Accuracies for the maximum of the SNP weights of the window (WssGBLUP_{Max}) and the sum of all SNP weights of the window (WssGBLUP_{Sum}) are presented for a window size of 40 consecutive SNP.

as WssGBLUP_{Max}. For FC, FU, RUA, and US, WssGBLUP_{Sum} was slightly more accurate (+1 percentage point) than WssGBLUP_{Max}. With WssGBLUP_{Sum}, accuracy was very similar to that obtained with WssGBLUP_{Max} whatever the breed and the trait.

In both the Alpine and Saanen breeds, WssGBLUP_{Sum} and WssGBLUP_{Max} were at least as accurate as, or more accurate than WssGBLUP, with differences ranging from +0 to +4 percentage points for RUA [WssGBLUP_{Max} compared with WssGBLUP in the Alpine breed ($P < 0.05$)] or FU [WssGBLUP_{Sum} compared with WssGBLUP in the Saanen breed ($P < 0.01$)].

Finally, on average for all traits, WssGBLUP_{Sum} and WssGBLUP_{Max} were significantly more accurate (+3 percentage points) than ssGBLUP in the Saanen breed, whereas WssGBLUP_{Sum} and WssGBLUP_{Max} were as accurate as ssGBLUP in the Alpine breed. In the Saanen breed, for traits with a QTL on chromosome 19, WssGBLUP_{Sum} and WssGBLUP_{Max} were significantly more accurate than ssGBLUP, from +3 percentage points to +7 percentage points for MY, PY, and UFP ($P < 0.01$). For RUA, improvement of accuracy with WssGBLUP_{Sum} and WssGBLUP_{Max} (+3 percentage points) was not significant. For the other traits, WssGBLUP_{Sum} and WssGBLUP_{Max} did not significantly outperform ssGBLUP (+0 to +2 percentage points). In the Alpine breed, no QTL was detected and the accuracy of all the methods was similar.

Our results are consistent with those reported by Zhang et al. (2016). We conclude that for polygenic traits, the same accuracy can be obtained with ssGBLUP, WssGBLUP_{Sum}, and WssGBLUP_{Max}, but genomic evaluations made with WssGBLUP are less accurate than with ssGBLUP. However, when QTL are detected for a trait, a slight gain in accuracy can be obtained with WssGBLUP_{Sum} and WssGBLUP_{Max} in

addition to that obtained with WssGBLUP compared with ssGBLUP.

CONCLUSIONS

Our aim was to investigate different genomic evaluation methods that allow information on the genetic architecture of traits (with or without QTL identified) to be incorporated. We compared ssGBLUP with weighted ssGBLUP and its alternatives (WssBLUP_{Sum} and WssGBLUP_{Max}) for the standard traits selected in the 2 main French dairy goat breeds. Weighted ssGBLUP is an iterative algorithm, and we confirmed that the highest accuracies were obtained at the second iteration. The weighted ssGBLUP and its alternatives were able to improve accuracy of genomic evaluations compared with ssGBLUP for traits with a QTL previously identified in a GWAS (MY, FY, PY, UFP, and RUA in the Saanen breed). Compared with ssGBLUP, the gain in accuracy was between 2 and 14% for weighted ssGBLUP and between 3 and 14% for WssGBLUP_{Sum} and WssGBLUP_{Max} in the Saanen breed. For traits with no identified QTL (FC, TA, FU, and US in the Saanen breed and all traits studied in the Alpine breed), Weighted ssGBLUP was less accurate than ssGBLUP (between -5 and 0%). With WssGBLUP_{Sum} and WssGBLUP_{Max}, the accuracy of the genomic evaluation was close to the accuracy achieved with ssGBLUP (between -2 and 4%). We will recommend the use of the WssGBLUP at the second iteration to predict GEBV of animals in French dairy goat breeding program.

ACKNOWLEDGMENTS

This study would not have been possible without the goat SNP50 BeadChip developed by the International

Goat Genome Consortium (IGGC; www.goatgenome.org). The authors thank Ignacy Misztal (University of Georgia, Athens) for the blup90iod2 program. The authors thank the French Genovicap and Phenofinlait programs [The French National Research Agency (ANR), Apis-Gene, specialized fund for agricultural and rural development (CASDAR), FranceAgriMer, France Genetique Elevage, and the French Ministry of Agriculture Agrifood, and Forestry] and the European 3SR project, which funded part of this work. The first author also received financial support from the Occitanie region and the French National Institute for Agricultural Research (INRA, France) Sélection Génomique (SELGEN) program (INCoMINGS).

REFERENCES

- Andonov, S., D. A. L. Lourenco, B. O. Fragomeni, Y. Masuda, I. Pocrnic, S. Tsuruta, and I. Misztal. 2017. Accuracy of breeding values in small genotyped populations using different sources of external information—A simulation study. *J. Dairy Sci.* 100:395–401. <https://doi.org/10.3168/jds.2016-11335>.
- Auvray, B., J. C. McEwan, S.-N. Newman, M. Lee, and K. G. Dodds. 2014. Genomic prediction of breeding values in the New Zealand sheep industry using a 50K SNP chip. *J. Anim. Sci.* 92:4375–4389. <https://doi.org/10.2527/jas.2014-7801>.
- Baloche, G., A. Legarra, G. Sallé, H. Larroque, J.-M. Astruc, C. Robert-Granié, and F. Barillet. 2014. Assessment of accuracy of genomic prediction for French Lacaune dairy sheep. *J. Dairy Sci.* 97:1107–1116. <https://doi.org/10.3168/jds.2013-7135>.
- Boichard, D., F. Guillaume, A. Baur, P. Croiseau, M. N. Rossignol, M. Y. Boscher, T. Druet, L. Genestout, J. J. Colleau, L. Journaux, V. Ducrocq, and S. Fritz. 2012. Genomic selection in French dairy cattle. *Anim. Prod. Sci.* 52:115–120. <https://doi.org/10.1071/AN11119>.
- Brito, L. F., S. M. Clarke, J. C. McEwan, S. P. Miller, N. K. Pickering, W. E. Bain, K. G. Dodds, M. Sargolzaei, and F. S. Schenkel. 2017. Prediction of genomic breeding values for growth, carcass and meat quality traits in a multi-breed sheep population using a HD SNP chip. *BMC Genet.* 18:7. <https://doi.org/10.1186/s12863-017-0476-8>.
- Carillier, C., H. Larroque, I. Palière, V. Clément, R. Rupp, and C. Robert-Granié. 2013. A first step toward genomic selection in the multi-breed French dairy goat population. *J. Dairy Sci.* 96:7294–7305. <https://doi.org/10.3168/jds.2013-6789>.
- Carillier, C., H. Larroque, and C. Robert-Granié. 2014. Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. *Genet. Sel. Evol.* 46:67. <https://doi.org/10.1186/s12711-014-0067-3>.
- Carillier-Jacquín, C., H. Larroque, and C. Robert-Granié. 2016. Including α s1 casein gene information in genomic evaluations of French dairy goats. *Genet. Sel. Evol.* 48:54. <https://doi.org/10.1186/s12711-016-0233-x>.
- Chen, C. Y., I. Misztal, I. Aguilar, S. Tsuruta, T. H. E. Meuwissen, S. E. Aggrey, T. Wing, and W. M. Muir. 2011. Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens. *J. Anim. Sci.* 89:23–28. <https://doi.org/10.2527/jas.2010-3071>.
- Christensen, O. F., P. Madsen, B. Nielsen, T. Ostensen, and G. Su. 2012. Single-step methods for genomic evaluation in pigs. *Animal* 6:1565–1571. <https://doi.org/10.1017/S1751731112000742>.
- Clément, V., D. Boichard, A. Piacère, A. Barbat, and E. Manfredi. 2002. Genetic evaluation of French goats for dairy and type traits. Pages 235–238 in *Proc. 7th World Congr. Genet. Appl. Livest. Prod.*, Montpellier, France.
- Duchemin, S. I., C. Colombani, A. Legarra, G. Baloche, H. Larroque, J.-M. Astruc, F. Barillet, C. Robert-Granié, and E. Manfredi. 2012. Genomic selection in the French Lacaune dairy sheep breed. *J. Dairy Sci.* 95:2723–2733. <https://doi.org/10.3168/jds.2011-4980>.
- Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: Application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal* 1:21–28. <https://doi.org/10.1017/S1751731107392628>.
- Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica* 136:245–257. <https://doi.org/10.1007/s10709-008-9308-0>.
- Grosclaude, F., M.-F. Mahé, G. Brignon, L. Di Stasio, and R. Jeunet. 1987. A Mendelian polymorphism underlying quantitative variations of goat α _{S1}-casein. *Genet. Sel. Evol.* 19:399–412. <https://doi.org/10.1186/1297-9686-19-4-399>.
- Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92:433–443. <https://doi.org/10.3168/jds.2008-1646>.
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92:4656–4663. <https://doi.org/10.3168/jds.2009-2061>.
- Legarra, A., and Z. G. Vitezica. 2015. Genetic evaluation with major genes and polygenic inheritance when some animals are not genotyped using gene content multiple-trait BLUP. *Genet. Sel. Evol.* 47:89. <https://doi.org/10.1186/s12711-015-0165-x>.
- Lourenco, D. A. L., I. Misztal, S. Tsuruta, I. Aguilar, E. Ezra, M. Ron, A. Shirak, and J. I. Weller. 2014. Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. *J. Dairy Sci.* 97:1742–1752. <https://doi.org/10.3168/jds.2013-6916>.
- Martin, P., I. Palière, C. Maroteau, P. Bardou, K. Canale-Tabet, J. Sarry, F. Woloszyn, J. Bertrand-Michel, I. Racke, H. Besir, R. Rupp, and G. Tosser-Klopp. 2017. A genome scan for milk production traits in dairy goats reveals two new mutations in Dgat1 reducing milk fat content. *Sci. Rep.* 7. <https://doi.org/10.1038/s41598-017-02052-0>.
- Martin, P., I. Palière, C. Maroteau, V. Clément, I. David, G. T. Klopp, and R. Rupp. 2018. Genome-wide association mapping for type and mammary health traits in French dairy goats identifies a pleiotropic region on chromosome 19 in the Saanen breed. *J. Dairy Sci.* 101:5214–5226. <https://doi.org/10.3168/jds.2017-13625>.
- Matilainen, K., M. Koivula, I. Strandén, G. P. Aamand, and E. A. Mäntysaari. 2016. Managing genetic groups in single-step genomic evaluations applied on female fertility traits in Nordic Red Dairy cattle. *Interbull Bull.* 50:71–75.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829.
- Misztal, I., S. E. Aggrey, and W. M. Muir. 2013. Experiences with a single-step genome evaluation. *Poult. Sci.* 92:2530–2534. <https://doi.org/10.3382/ps.2012-02739>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs. *Commun. No. 28–07. Proc. 7th World Congr. Genet. Appl. Livest. Prod.*, Montpellier, France.
- Mucha, S., R. Mrode, M. Coffey, M. Kizilaslan, S. Desire, and J. Conington. 2018. Genome-wide association study of conformation and milk yield in mixed-breed dairy goats. *J. Dairy Sci.* 101:2213–2225. <https://doi.org/10.3168/jds.2017-12919>.
- Mucha, S., R. Mrode, I. MacLaren-Lee, M. Coffey, and J. Conington. 2015. Estimation of genomic breeding values for milk yield in UK dairy goats. *J. Dairy Sci.* 98:8201–8208. <https://doi.org/10.3168/jds.2015-9682>.
- Onogi, A., A. Ogino, T. Komatsu, N. Shoji, K. Shimizu, K. Kurogi, T. Yasumori, K. Togashi, and H. Iwata. 2015. Whole-genome prediction of fatty acid composition in meat of Japanese Black cattle. *Anim. Genet.* 46:557–559. <https://doi.org/10.1111/age.12300>.
- Teissier, M., H. Larroque, and C. Robert-Granié. 2018. Weighted single-step genomic BLUP improves accuracy of genomic breeding values for protein content in French dairy goats: A quantitative

- trait influenced by a major gene. *Genet. Sel. Evol.* 50:31. <https://doi.org/10.1186/s12711-018-0400-3>.
- Tosser-Klopp, G., P. Bardou, O. Bouchez, C. Cabau, R. Crooijmans, Y. Dong, C. Donnadiou-Tonon, A. Eggen, H. C. M. Heuven, S. Jamli, A. J. Jiken, C. Klopp, C. T. Lawley, J. McEwan, P. Martin, C. R. Moreno, P. Mulsant, I. Nabihoudine, E. Pailhoux, I. Palhière, R. Rupp, J. Sarry, B. L. Sayre, A. Tircazes, J. Wang, W. Wang, and W. Zhang. International Goat Genome Consortium. 2014. Design and Characterization of a 52K SNP chip for goats. *PLoS One* 9:e86227. <https://doi.org/10.1371/journal.pone.0086227>.
- Van Sickle, J. 2003. Analyzing correlations between stream and watershed attributes. *J. Am. Water Resour. Assoc.* 39:717–726. <https://doi.org/10.1111/j.1752-1688.2003.tb03687.x>.
- Venot, E., D. Boichard, V. Ducrocq, S. Fritz, H. Larroque, F. Tortereau, J.-M. Astruc, A. Barbat, M. Barbat, A. Baur, P. Boulesteix, C. Carillier-Jacquín, P. Croiseau, M.-N. Fouilloux, A. Gion, C. Hoze, A. Launay, R. Lefebvre, A. Legarra, V. Loywyck, I. Palhière, F. Phocas, J. Promp, C. Robert-Granie, R. Rupp, R. Saintilan, M.-P. Sanchez, T. Tribout, A. Vinet, and S. Mattalia. 2017. French genomic experience: Genomics for all ruminant species. In 40th Biennial Session of ICAR.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* 94:73–83. <https://doi.org/10.1017/S0016672312000274>.
- Wolc, A., A. Kranis, J. Arango, P. Settar, J. E. Fulton, N. P. O'Sullivan, A. Avendano, K. A. Watson, J. M. Hickey, G. de los Campos, R. L. Fernando, D. J. Garrick, and J. C. M. Dekkers. 2016. Implementation of genomic selection in the poultry industry. *Anim. Front.* 6:23–31. <https://doi.org/10.2527/af.2016-0004>.
- Zhang, Z., X. Ding, J. Liu, D.-J. de Koning, and Q. Zhang. 2011. Genomic selection for QTL-MAS data using a trait-specific relationship matrix. *BMC Proc.* 5:S15. <https://doi.org/10.1186/1753-6561-5-S3-S15>.
- Zhang, Z., M. Erbe, J. He, U. Ober, N. Gao, H. Zhang, H. Simianer, and J. Li. 2015. Accuracy of whole-genome prediction using a genetic architecture-enhanced variance-covariance matrix. *G3 Genes Genomes Genet.* 5:615–627. <https://doi.org/10.1534/g3.114.016261>.
- Zhang, X., D. Lourenco, I. Aguilar, A. Legarra, and I. Misztal. 2016. Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. *Front. Genet.* 7:151. <https://doi.org/10.3389/fgene.2016.00151>.
- Zhou, L., R. Mrode, S. Zhang, Q. Zhang, B. Li, and J.-F. Liu. 2018. Factors affecting GEBV accuracy with single-step Bayesian models. *Heredity* 120:100–109. <https://doi.org/10.1038/s41437-017-0010-9>.