Validation of single-step genomic BLUP random regression test-day models and SNP effects analysis on milk yield in French Saanen goats

M. Arnal, C. Robert-Granié, V. Ducrocq, H. Larroque

J. Dairy Sci. 106:4813–4824
https://doi.org/10.3168/jds.2022-22550

Validation of single-step genomic BLUP random regression test-day models and SNP effects analysis on milk yield in French Saanen goats

M. Arnal,[1,2]* C. Robert-Granié,[1] V. Ducrocq,[3] and H. Larroque[1]
[1]GenPhySE, Université de Toulouse, INRAE, INPT, ENVT, 31326 Castanet-Tolosan, France
[2]Institut de l'Elevage, Chemin de Borde Rouge, 31326 Castanet-Tolosan Cedex, France
[3]Université Paris-Saclay, INRAE, AgroParisTech, UMR GABI, 78350 Jouy-en-Josas, France

ABSTRACT

The shape of the lactation curve is linked to an animal's health, feed requirements, and milk production throughout the year. Random regression models (RRM) are widely used for genetic evaluation of total milk production throughout the lactation and for milk yield persistency. Genomic information used with the single-step genomic BLUP method (ssGBLUP) substantially improves the accuracy of genomic prediction of breeding values in the main dairy cattle breeds. The aim of this study was to implement an RRM using ssGBLUP for milk yield in Saanen dairy goats in France. The data set consisted of 7,904,246 test-day records from 1,308,307 lactations of Saanen goats collected in France between 2000 and 2017. The performance of this type of evaluation was assessed by applying a validation step with data targeting candidate bucks. The model was compared with a nongenomic evaluation and a traditional evaluation that use cumulated performance throughout the lactation model (LM). The incorporation of genomic information increased correlations between daughter yield deviations (DYD) and estimated breeding values (EBV) obtained with a partial data set for candidate bucks. The LM and the RRM had similar correlation between DYD and EBV. However, the RRM reduced overestimation of EBV and improved the slope of the regression of DYD on EBV obtained at birth. This study shows that a genomic evaluation from a ssGBLUP RRM is possible in dairy goats in France and that RRM performance is comparable to a LM but with the additional benefit of a genomic evaluation of persistency. Variance of adjacent SNPs was studied with LM and RRM following the ssGBLUP. Both approaches converged on approximately the same regions explaining more than 1% of total variance. Regions associated with persistency were also found.

Received July 19, 2022.
Accepted January 4, 2023.
*Corresponding author: mathieu.arnal@idele.fr

Key words: random regression model, single-step genomic BLUP, dairy goat persistency

INTRODUCTION

Saanen is one of the most common dairy goat breeds in the world (Currò et al., 2019) and one of the 2 major dairy goat breeds in France. In France, genetic evaluation for milk yield in dairy goats has traditionally been based on cumulative milk production over 250 d [i.e., on a lactation model (**LM**); Clément et al., 2002], and interpolation between test-day records (**TD**) has been based on the so-called Fleischmann method (Sargent et al., 1968).

Many countries use random regression models (**RRM**) to estimate genetic breeding values of dairy animals based on their TD (Oliveira et al., 2019a). The RRM can better account for environmental effects to compute EBV for lactation milk yield (Druet et al., 2003). The RRM also represents an alternative model for generating EBV for milk yield persistency (Oliveira et al., 2019d). Persistency is commonly defined as the rate of decline in milk production after its peak (Cole and Null, 2009). Persistency EBV can be obtained from the eigenvectors of the genetic (co)variance matrices of RRM; the first eigenvector is almost constant throughout lactation, but the second eigenvector can be made negative at the beginning of lactation, null in the middle of lactation, and positive at the end of lactation, as presented in Druet et al. (2003). These eigenvectors can be used instead of Legendre polynomials or other alternatives to model the genetic and permanent environment effects as follows: the first EBV obtained with the first eigenvector can be interpreted as the EBV of level of production throughout the lactation, and the second EBV obtained with second eigenvector measures the persistency EBV. These EBVs are also of interest because the genetic correlation between these 2 traits is null by construction. In a breeding objective with milk level and milk persistency, a null correlation between those is a desirable feature. The persistency EBV obtained with the second eigenvector can be used to select animals which produce less milk at the beginning of the lactation cycle and more at the end than an average animal that produces the same amount of milk throughout the lactation. This selection strategy may be useful to reduce health disorders by limiting milk production at the peak of the lactation. Moreover, it allows to better spread milk production throughout the year, which is an advantage in dairy goats for which production follows strong seasonal patterns with most kidding happening in the middle of winter. The main disadvantage of RRMs is that they are time-consuming compared with LMs (Schaeffer et al., 2000).

Genomic information from SNP is commonly used in genetic evaluation to improve the accuracy of EBV (R2D2 Consortium et al., 2021). Single-step genomic BLUP (**ssGBLUP**) is the reference method for estimating genetic merit values in populations in which not all animals are genotyped (Legarra et al., 2009; Aguilar et al., 2010). It directly uses the phenotypes, pedigree and SNP genotype information to build a relationship matrix (**H**) associating genotyped animals with nongenotyped animals (Legarra et al., 2009). Building **H** can be quite time-consuming when there are many genotyped animals, so the **H** matrix does not often get used in RRM despite all the advantages of this type of model (Oliveira et al., 2019a).

The validation of a genetic evaluation usually uses a cohort of newborn male candidates for an artificial insemination scheme as the target population. Indeed, it is important to select the animals that have the best genetic values when they are newborn to save breeding costs. In a genetic evaluation validation, the EBV of bucks obtained with a partial data set (after deletion of recent data to mimic the absence of information in the newest generation) are compared with their daughter yield deviation (**DYD**) obtained with the complete data set (VanRaden and Wiggans, 1991). Two parameters are usually considered: the dispersion of EBV, which is measured through the regression coefficient of the real measure of daughter production (DYD) on the EBV obtained with the partial data set (this regression slope should be close to 1), and a measure of bias defined as the difference between the DYD means and the EBV obtained from the partial data set (a value close to zero is expected).

After a ssGBLUP evaluation, an association study between SNPs and traits studied (Aguilar et al., 2014) can be performed to study the genomic regions associated with the trait. Several studies have used an RRM to investigate the regions associated with persistency in cattle (Strucken et al., 2012; Oliveira et al., 2019c), and Cardona et al. (2016) showed in goats that genes can be expressed differently throughout the lactation cycle. The variance explained by n adjacent SNPs (for example $n = 10$) can be calculated to reduce noise and obtain a better signal than when considering SNPs individually (Wang et al., 2014).

The aim of this study was to implement a genomic ssGBLUP RRM evaluation in France for Saanen dairy goats and to evaluate the benefits of this type of model in our French population. To do this, we evaluated the predictive power of genetic values at birth of a cohort of young males (measured by accuracy, dispersion, and bias) from either RRM and LM with and without genomic information. Then, we compared analyses of SNP effects using an LM and an RRM, and we studied genomic regions associated with persistency evaluated with an RRM.

MATERIALS AND METHODS

Data

The study was based on already available data; therefore, ethical approval was not required.

Only the first 3 parities of Saanen goats were analyzed. Each lactation included TD between DIM 7 and 270. The trait analyzed was milk yield (**MY**). To mimic a routine evaluation, all the lactations (from first to third parity) retained for the official evaluation (Larroque et al., 2011) were kept. Goats were milked twice a day, and one or 2 milkings were measured; if only a single daily measurement (32% of our data) was available, the production was multiplied by a coefficient that considers the difference in production between morning and evening milkings. This coefficient was determined from the quantity of milk measured during the TD and the quantity of milk in the cooling tank that cumulates production over several milkings. The use of this coefficient allows the inclusion in the study of the records of protocols T (ICAR, 2018), where production is measured alternately in the morning or evening. It is a refinement compared with a simple multiplication by 2. The total milk yields throughout the lactation were calculated in the same way as for routine genetic evaluations, using the Fleischmann method (Sargent et al., 1968). The Fleischmann method calculates total production by adding the cumulative production of the different periods defined by the TD (ICAR, 2022).

To test the quality of prediction of the genetic evaluation, a situation close to the current breeding scheme was assumed: the future bucks are selected as newborns. The newborn bucks, called candidate bucks, were the bucks born between 2010 and 2013. So, the complete phenotypic data set was split up and the part of the data collected from 2000 to 2011 formed the partial data set. The complete data set (c) consisted

Table 1. Test day (TD) and lactation numbers stratified by parity for the different data sets

Data set	Parity	TD	Lactation
Complete data set (2000–2017)	First	3,466,430	570,641
	Second	2,628,369	431,407
	Third	1,809,447	304,656
	Total	7,904,246	1,306,704
Partial data set (2000–2011)	First	2,566,171	413,817
	Second	1,942,222	311,255
	Third	1,346,841	221,250
	Total	5,855,234	946,322

of 7,908,192 TD records from 1,308,307 lactations of Saanen goats, collected in France between 2000 and 2017. Table 1 gives details on the number of TD and of lactations per data set. The pedigree consisted in 818,702 animals (40% of French dairy goats have a known sire and dam).

Goats were genotyped with the Illumina goat SNP50 BeadChip (Tosser-Klopp et al., 2014). The rules applied for SNP quality control were the same as in Teissier et al. (2019). At the end of quality control, 47,206 SNPs were kept. Table 2 shows the number of genotyped animals and the number of animals in the pedigree: 1,242 Saanen genotyped animals (430 Saanen genotyped bucks) were included in the evaluation. There were 133 newborn genotyped candidate bucks in the Saanen pedigree (Table 2). The number of candidates was not large, but adding more candidates would have reduced further the reference population.

Models of Evaluation

Four different evaluations were implemented and compared. The evaluations were run using the blup90iod2 software (Misztal et al., 2002).

Classical LM

A LM close to the one routinely used for official genetic evaluations in France (Larroque et al., 2011) was implemented. The genetic evaluation model based on these phenotypes was

$$y_{rnjlkmpi} = H_{rji} + A_{jkp} + M_{rjlp} + D_{rjmp} + a_n + p_n + e_{rnjlkmpi},$$

where $y_{rnjlkmpi}$ is the observed lactation of goat n, in production year j (2000, …, 2017), in parity r (1, 2, 3), belonging to kidding age, class k (7 classes for first lactation, in months: 9–11, 12, 13, 14, 15, 16, +17; 6 classes for second lactation: 21–23, 24, 25, 26, 27, +28; 6 classes for third lactation: 31–35, 36, 37, 38, 39, +40),

Table 2. Numbers of genotyped animals, with reference and candidate bucks stratified by year of birth, and number of individuals in the pedigree file (pedigree, n = 818,702; genotyped female, n = 812)

Genotyped animal	Year of birth	No. of bucks	Total
Reference bucks	1998	19	297
	1999	21	
	2000	18	
	2001	25	
	2002	28	
	2003	29	
	2004	25	
	2005	29	
	2006	25	
	2007	29	
	2008	21	
	2009	28	
	2010	30	
Candidate bucks	2011	32	133
	2012	35	
	2013	36	

kidding month, class l (7 classes: January, February, March–May, June–September, October, November, December), dry period length, class m (6 classes, in days: [First parity], [0,50], [50,75], [75,100], [100,125], 125+), in region p (4 classes: North-West, North-East, South-West, South-East) and herd i. H_{rji} is the fixed effect of herd; A_{jkp} is the fixed effect of age at kidding; M_{rjlp} is the fixed effect of kidding period; D_{rjmp} the fixed effect of length of the dry period; a_n is the additive genetic breeding value that followed a normal distribution (mean: $\mu_{\mathrm{LACT}} = 0$, variance $= \sigma_{\mathrm{LACT}}{}^2\mathbf{A}$). $\mathbf{A}$ is the additive genetic relationship matrix based on pedigree information. $\mathbf{H}$ is the additive genetic relationship matrix based on pedigree and genomic information, as in Legarra et al. (2009) with

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau\left[(1-w)\mathbf{G} + w\mathbf{A}_{22}\right]^{-1} - \omega\mathbf{A}_{22}^{-1} \end{bmatrix}.$$

We used $\tau = 1$, $w = 0.05$, and $\omega = 1$, which are the default values used in blup90iod2 (Misztal et al., 2002). For the scaling of $\mathbf{G}$, the mean of the diagonal of $\mathbf{G}$ was set equal to the mean of the diagonal of $\mathbf{A}$. The mean of the off-diagonal elements of $\mathbf{G}$ was set to be equal of the off-diagonal elements of $\mathbf{A}$. p_n is the permanent environment value that followed a normal distribution (mean: $\mu_{\mathrm{PE\_LACT}} = 0$, variance: $\sigma_{\mathrm{PE\_LACT}}{}^2$), and $e_{rnjlkmpi}$ is the residual term.. To account for missing ancestral pedigree, we used metafounders (**MF**) as in Legarra et al. (2015). For the relationship matrix $\boldsymbol{\Gamma}$ among **MF** defined in Garcia-Baccino et al. (2017), we used a modified $\boldsymbol{\Gamma}$, called the gamma-robust estimator. This matrix is based on the median element of the original $\boldsymbol{\Gamma}$ calculated with the gammaf90 software.

Using the partial data set, the EBV obtained from the model that used pedigree information to build the relationship matrix were called **A_LM**, and the GEBV obtained from the ssGBLUP model were called **H_LM**.

Random Regression Model (RRM)

Various preliminary studies were performed (Arnal et al., 2019, 2020) to determine which type of functions should be used in the RRM. Based on these studies, the model used here was the "EGV_PM" model described in Arnal et al. (2020).

$$y_{rsijklmdgnp} = HTD_{ri} + A_{jkp} + M_{rjlp} + D_{rjmp} + \sum_{o=1}^{6}\theta_{ko}N_{(o,d)} + \sum_{o=1}^{6}\tau_{rlo}N_{(o,d)} + \sum_{o=1}^{6}\pi_{rmo}N_{(o,d)} + \sum_{o=1}^{4}\gamma_{so}M_{(o,g)} + \sum_{o=1}^{2}b_{son}\chi_{(o,d)} + \sum_{o=1}^{2}c_{son}\chi_{(o,d)} + e_{rsijklmdgnp},$$

where $y_{rsijklmdgnp}$ is the observed DIM d (7, . . ., 270) of goat n after g d of gestation (0, . . ., 100), in production year j (2000, . . ., 2017), in parity r (1, 2, 3), in parity, class s [primiparous (P) or multiparous (M)], belonging to kidding age, class k (7 classes for first lactation, in months: 9–11, 12, 13, 14, 15, 16, +17; 6 classes for second lactation: 21–23, 24, 25, 26, 27, +28; 6 classes for third lactation: 31–35, 36, 37, 38, 39, +40), kidding month, class l (7 classes: January, February, March–May, June–September, October, November, December), dry period length, class m [6 classes, in days: (first parity), (0,50), (50,75), (75,100), (100,125), 125+], in region p (4 classes: North-West, North-East, South-West, South-East), herd × test-date class i. HTD_{ri} is the fixed effect of herd test-date; A_{jkp} is the fixed effect of age at kidding; M_{rjlp} is the fixed effect of kidding period; D_{rjmp} is the fixed effect of length of the dry period; θ_{ko}, $\theta\tau_{rlo}$, π_{rmo}, γ_{so} are fixed regression coefficients for age at kidding, kidding month, dry period length, and gestation stage, respectively; $N_{(o,d)}$ is the oth covariate at time d of a cubic natural spline function with 6 knots at d = 7, 20, 50, 110, 190, 270; $M_{(o,g)}$ is the oth covariate at time d of a cubic natural spline function with 4 knots at d = 31, 53, 76, 100 (between g = 0 and g = 30, the coefficients were assumed to be equal to 0; if gestation stage was greater than 100, it was rounded down to 100); b_{son} and c_{son} are the random additive genetic and permanent environmental regression coefficients for the oth eigenvectors of the genetic (co)variances matrix obtained with a second-order Legendre polynomial model reduced to rank 2, $\chi_{(o,c)}$ is the value of the oth eigenvector of the genetic (co)variances matrix obtained with a second-order Legendre polynomial model reduced to rank 2 at DIM d for parity class s, and $e_{rsijklmdgnp}$ is the residual term. Due to software limitations, we first considered a homogeneous residual variance. The random additive genetic regression coefficient for the first eigenvector, b_1, is noted ***LEV*** (related to production level), and the second eigenvector, b_2, is noted ***PERS*** (for persistency), as presented in Druet et al. (2005); LEV is noted **LEV_P** for primiparous, or **LEV_M** for multiparous; PERS is noted **PERS_P** for primiparous or **PERS_M** for multiparous. The second and third parities were considered separately in the fixed part to be more precise because phenotypically the second and third lactation are different in the shape of lactation curve and in the level of production. However, we kept them together for the genetic part because the genetic correlation between them was close to 1 (Arnal et al., 2020) and keeping them together substantially reduces the size and complexity of the model.

The EBVs for total production were computed separately for primiparous (**primi**) and multiparous (**multi**) goats. For that purpose, each daily EBV was computed as

$$EBV_{primi,d} = \text{LEV\_P} \times \chi_{(1,d)} + \text{PERS\_P} \times \chi_{(2,d)},$$

$$EBV_{multi,d} = \text{LEV\_M} \times \chi_{(3,d)} + \text{PERS\_M} \times \chi_{(4,d)}.$$

Daily EBV were then added to obtain a total EBV for primiparous (EBV_{primi}) and multiparous (EBV_{multi}) animals to have a kilograms of MY as the common unit and be comparable (on the same scale) with an EBV from a LM as in the following:

$$EBV_{primi} = \sum_{d=1}^{264} EBV_{primi,d}$$

$$EBV_{multi} = \sum_{d=1}^{264} EBV_{multi,d.}$$

Finally, EBV from primiparous and multiparous animals were combined as follows to have a readily comparable basis with the LM that has a unique EBV for the 3 lactations:

$$EBV_{tot} = 0.33 \times EBV_{primi} + 0.66 \times EBV_{multi}.$$

The 2 coefficients were chosen to give a same weight to the first 3 parities.

The same procedure was applied to obtain an EBV for milk production level and persistency.

$$lev_{tot} = 0.33 \times lev_{primi} + 0.66 \times lev_{multi},$$

$$pers_{tot} = 0.33 \times pers_{primi} + 0.66 \times pers_{multi}.$$

Using the partial data set, EBV_{tot} from the model that used only pedigree information to build the relationship matrix was called **A_RRM**, $GEBV_{tot}$ from the ssGBLUP model was called **H_RRM**, lev_{tot} from the model that used only pedigree information to build the relationship matrix was called **A_LEV**, lev_{tot} from the model that used genomic information in addition to pedigree to build the relationship matrix as in Legarra et al. (2009) was called **H_LEV**, $pers_{tot}$ from the model that used only pedigree information to build the relationship matrix was called **A_PERS**, and $pers_{tot}$ from the ssGBLUP model was called **H_PERS**.

Daughter Yield Deviation

The DYD was calculated from an RRM using the complete data set but without genomic information. The DYD term denotes total lactation production obtained as EBV_{tot}; **DYD_LEV** and **DYD_PERS** are the DYD for LEV and PERS calculated as lev_{tot} and $pers_{tot}$.

The DYD were calculated using the *genekit* software (Ducrocq, 1998), as in Täubert et al. (2010).

Criteria for Comparing Evaluations

The number of progeny of candidate bucks in 2018 was observed and the *genekit* software was used to estimate the reliability obtained in 2018, as in Ducrocq and Schneider (2007). Because the number of multiparous progeny differed strongly between bucks (from 6 to 152, with a median of 40 in the Saanen breed), the calculations of bias, slopes, and correlations were weighed according to the number of multiparous progeny.

Bias

The means of DYD and EBV were set to zero for the bucks (without restriction to AI bucks) that had more than 25 multiparous progeny born in 1997. The bias was calculated as

$$\mu_{DYD,p} = \frac{\overline{\hat{u}_p - DYD}}{\sigma_{DYD}},$$

where μ is bias, $\hat{u}_p$ is EBV obtained with partial data, and σ_{DYD} is a standard deviation of DYD.

Slope

The slope was calculated as

$$b_{DYD,p} = \frac{cov\left(\hat{u}_p, DYD\right)}{var\left(\hat{u}_p\right)},$$

where $b_{DYD,p}$ is the regression slope of DYD on EBV obtained with the partial data set.

Correlation

The correlation was calculated as

$$\rho_{DYD,p} = \frac{cov\left(\hat{u}_p, DYD\right)}{\sqrt{var\left(DYD\right) var\left(\hat{u}_p\right)}},$$

where $\rho_{DYD,p}$ is the correlation coefficient between EBV obtained with the partial data set and the DYD.

Analysis of SNPs' Effects

The variance explained by 10 adjacent SNPs was calculated after the ssGBLUP of H_LM and H_RRM with the complete data set using POSTGSF90 software (Aguilar et al., 2014). The sum of variance percentages is not equal to 100% of the total variance because the segments were overlapping.

RESULTS

Correlations Between Evaluations

The correlations between EBV for candidate bucks using the partial data set (without progeny performance) are presented in Table 3. The correlation between A_LM and H_LM were high (0.82). The correlation between A_RRM and H_RRM were very close to the correlations between A_LM and H_LM (0.81), which means that the genomic information brought the same changes in both models (LM or RRM). The correlation between A_LM and A_RRM were equal to 0.95. The correlation between H_LM and H_RRM were close to the correlation between A_RRM and H_RRM (0.96). The biggest changes were obtained when adding genomic information and not by changing the type of model (LM and RRM).

Table 3. Correlations between EBV from the different models of the candidate bucks calculated with the partial data set[1]

Model	A_LM	H_LM	A_RRM
H_LM	0.82		
A_RRM	0.95	0.76	
H_RRM	0.80	0.96	0.81

[1]A_LM: lactation model without genomic information used to build the relationship matrix. H_LM: lactation model with genomic information used to build the relationship matrix. A_RRM: random regression model without genomic information used to build the relationship matrix. H_RRM: random regression model with genomic information used to build the relationship matrix. Correlations are progeny-number-weighed.

Validation of the Different Genetic Evaluations

We used various criteria to test the predictive abilities of the different models, comparing the EBV of newborn genotyped bucks to their DYD obtained from the performances of their progeny. A model is desirable if the EBV of the newborn bucks are close to their mean DYD (studied by the bias), with a correlation close to one (assessed through the accuracy criterion) and a slope of the regression of the DYD on the EBV close to one, which is a measure of closeness and dispersion of the EBV. The results for LEV were the same as the results for RRM ($EBVtot$) observed in a previous study (Arnal et al., 2019), with a correlation close to 1 between A_RRM and A_LEV.

Correlations

Correlations between the DYD from complete data set and A_LM, A_RRM, H_LM and H_RRM obtained with partial data set are presented in Figure 1. The correlations between DYD and H_LM was 0.41 and between DYD and H_RRM was 0.43. The use of genomic information improved the correlations by 0.09 in the LM and by 0.11 in RRM. Correlations between DYD and LM or RRM were very close with or without genomic information. Correlations between DYD for PERS and the EBV for PERS obtained with partial data set with or without genomic information are presented in Figure 2. The correlation for persistency was 0.46 without genomic information and 0.52 with genomic information. Correlations for persistency were higher than correlations for production level. The genomic information improved LEV correlations more than PERS one (around +0.06).

Slope

Slopes were inferior to 1, indicating an overdispersion of buck EBV at birth. The RRM gave slopes closer to 1 than the LM without genomic information (0.6 for A_LM and 0.68 for A_RRM) and with genomic information (0.69 for H_LM and 0.86 for H_RRM; Figure 1). The use of genomic information improved the slopes with both models. The slope was better for PERS than EBV_{tot} without genomic information (0.73) but the use of genomic information led to less increase in the slope of PERS compared with EBV_{tot} (+0.04 vs. 0.17; Figure 2).

Bias

The LM introduced substantial bias, at 0.12 standard deviation on DYD. The bias was always positive, which means that EBV_{tot} values were overestimated. The RRM introduced less bias than the LM (Figure 1), at 0.06 standard deviations on DYD. The use of genomic information did not affect the bias. The means of A_PERS and H_PERS were smaller than DYD_PERS (−0.04; Figure 2). The bias for PERS was negative, which means that the EBV of PERS were underestimated compared with DYD. The use of genomic information had no effect on the bias, as for EBV_{tot}.

Evolution of Mean EBV Over the Years

Averages of EBV by birth year of bucks from the partial data calculated by RRM and LM were plotted in Figure 3. The means of DYD of bucks by birth years were also plotted. The EBV from the partial data set were used to calculate genetic trends in EBV to be compared against the genetic trends in DYD. The EBV from LM and RRM with or without genomic information followed the same trend pattern as DYD, particularly for bucks with progeny (bucks born before 2010). However, the means EBV from LM and RRM with or without genomic information by year of birth were not the same as DYD means by year of birth as it could be seen for the bias. The RRM EBV means were closer to DYD means than LM EBV. The mean differences increased from 1998, with a difference of 10 kg between RRM EBV and DYD (20 kg between LM EBV and DYD means), which was close to 40 kg in 2013 for RRM (80 kg between LM EBV and DYD means). Similar to A_RRM and H_RRM, A_LM and H_LM were very close to each other. For persistency, the evolution of the average persistence of EBV and DYD by year of birth was the same, even for bucks without offspring born between 2010 and 2013.

ANOVA Explained by Adjacent SNPs

Figure 4 shows the percentage of variance represented by overlapping segments of 10 SNPs for chromosomes 6 and 19. These 2 chromosomes contained segments repre-

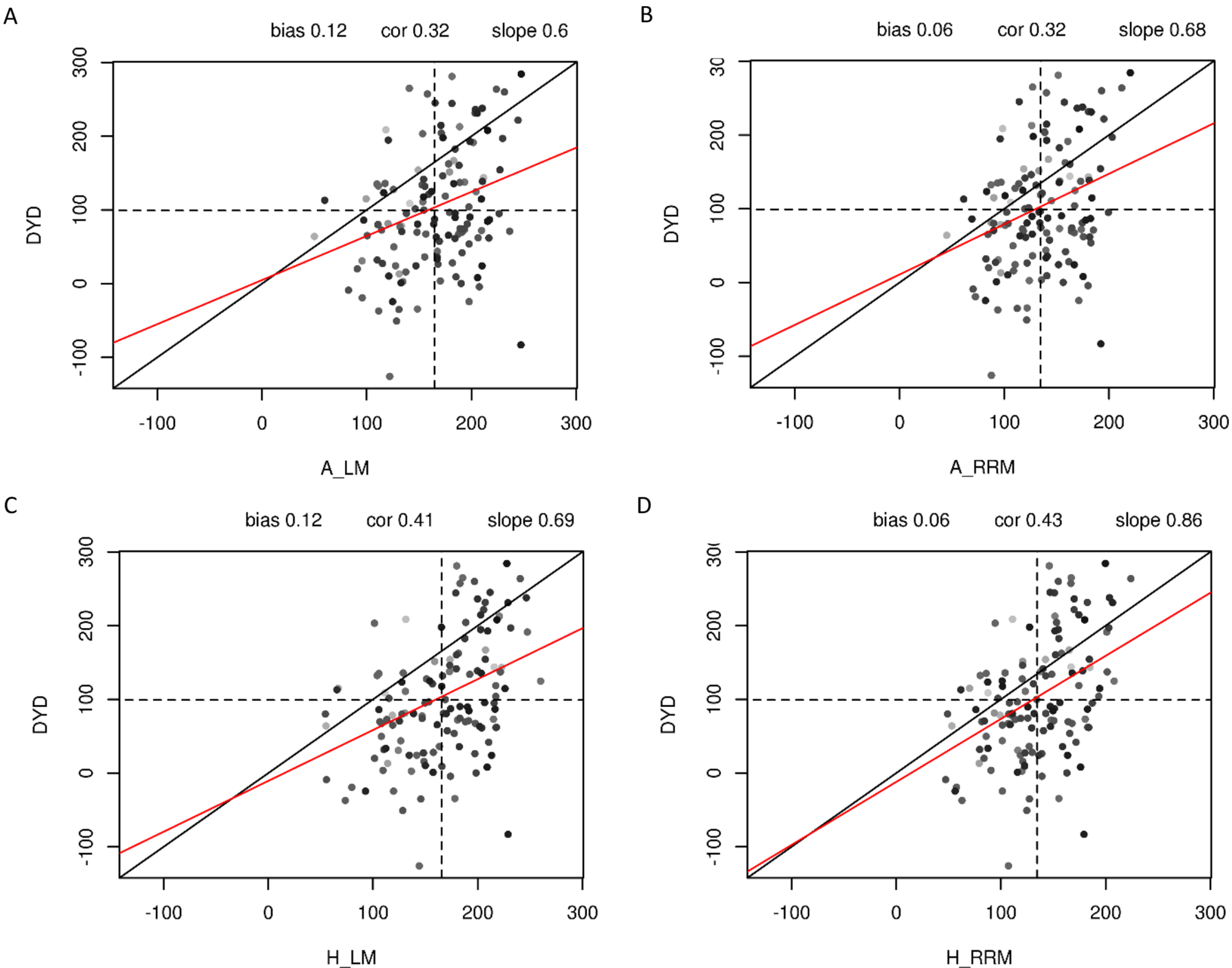

Figure 1. Plot between the buck-candidate EBV from the partial data set from A_LM (A), A_RRM (B), H_LM (C), H_RRM (D), and their daughter yield deviations (DYD). A_LM: lactation model without genomic information used to build the relationship matrix. H_LM: lactation model with genomic information used to build the relationship matrix. A_RRM: random regression model without genomic information used to build the relationship matrix. H_RRM: random regression model with genomic information used to build the relationship matrix. In black, the line y = x; in red, the line y = ax + b; the horizontal line is y = mean DYD; the vertical line is x = mean EBV total. cor = correlation between x and y.

senting more than 1% of the total variance. Percentages of variance explained by 10 adjacent SNPs segments representing more than 1% of variance are presented in Supplemental Table S1 (https://figshare.com/articles/figure/JDS_ARNAL_SUPPLEMENTARY_MATERIAL_TABLE_S1/22592416; Arnal et al., 2023). On chromosome 6, one segment (85.9–86.0 Mb) at the region of caseins (CSN1S1, CSN2), represented more than 1% of variance in LM (1.4%) and LEV_M (1.9%). On chromosome 19, one region (25.6–29.1 Mb) had several segments representing more than 1% of total variance for LM, LEV_P, and LEV_M. A segment between 26.1 and 26.6 Mb in the ALOX12 gene region explained the maximum of variance for 3 traits, with 3.3% for LM, 3.9% for LEV_P, and 2.9% for LEV_M. This segment represented 0.9% of variance for PERS_P. In both chromosomes, LM had an intermediate percentage between LEV_P and LEV_M, which is logical because it is composed of LEV_P and LEV_M EBV.

DISCUSSION

Validation of the Different Genetic Evaluation Models

The DYD from RRM were considered as reference values rather than DYD from LM, because Arnal et

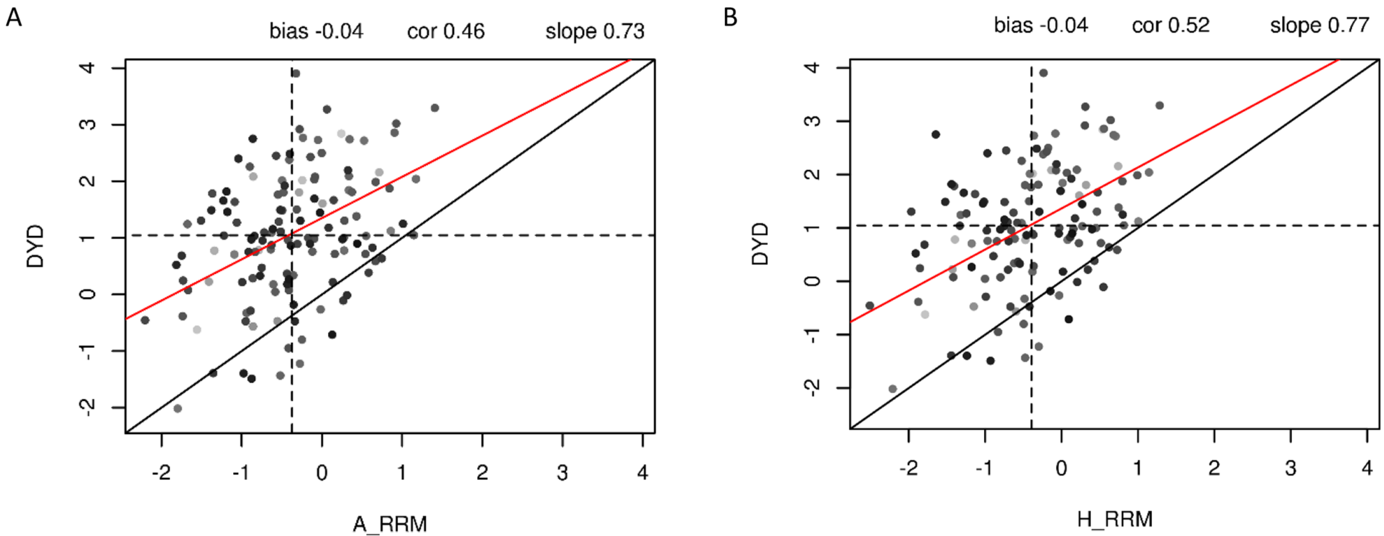

Figure 2. Plot between the buck-candidate EBV for persistency from the partial data set from A_RRM (A), H_RRM (B) and their daughter yield deviations (DYD). A_RRM: random regression model without genomic information used to build the relationship matrix. H_RRM: random regression model with genomic information used to build the relationship matrix. In black, the line y = x; in red, the line y = ax + b; the horizontal line is y = mean DYD; the vertical line is x = mean EBV total. cor = correlation between x and y.

al. (2019) showed that modeling the genetic and permanent environment effects according to DIM is more accurate with RRM than with LM. Moreover, with our RRM, other environmental parameters in the fixed part of the models were more precisely defined (gestation stage, age at kidding, month of kidding and dry period length depending on lactation stage, herd TD effect instead of herd-year effect, and separation of parities). The main difference between LM and RRM was in the means of EBV or DYD. Other studies (Arnal et al. 2019, 2020) used a heterogeneous residual variance for DIM and found that considering residual variance as heterogeneous preformed somewhat better for convergence properties. However, using a heterogeneous residual variance is a more complex task that cannot be done with the software (blup90iod2; Misztal et al., 2002) used here. For total production throughout the lactation with A or H and LM or RRM, the EBV were overestimated, with overdispersion and poor accuracy (validation correlations <0.52). These results are consistent with those obtained by Kang et al. (2017) in their simulation study using an RRM with an **A** and **H** matrix. In our study, the bias with RRM was reduced compared with LM. The lower bias and better slope with RRM were probably a consequence of a better fit of the model to the data, as described by Schaeffer and Jamrozik (2008). Here, regression slopes were less than 1, so the EBV were overdispersed, as in the majority of genomic studies (Legarra and Reverter, 2018). For the construction of the **H** matrix, the parameters ω and τ used were default parameters, with $\omega = 1$ and $\tau = 1$ set based on recommendations in other studies (Kang et al., 2018; Oliveira et al., 2019d). Misztal et al. (2017) showed the importance of these values on bias. In dairy cattle, Oliveira et al. (2019b) found, as in our study, that RRM do not improve the validation correlations compared with LM using **A** or **H**. The LM does not model the shape of the lactation curve but relies on a phenotype that takes it into account. This was certainly one of the reasons why the LM and the RRM were close in terms of correlation. Correlations between EBV of genotyped bucks and DYD were higher in both the LM and RRM when using genomic information for persistency. Several studies (Koivula et al., 2015; Mucha et al., 2015; Baba et al., 2017; Kang et al., 2017, 2018; Oliveira et al., 2019b) that have compared A_RRM with H_RRM found an increase in validation correlations and a slope closer to 1, which was due to **H** considering the Mendelian sampling term. Oliveira et al. (2019b) obtained similar results when they compared A_LM with H_LM.

The reference population in this study was very small, which limited the performance of genomic evaluation. With more animals genotyped, correlations and slope should improve. However, this reference Saanen population is very convenient, because it features a large share of AI bucks. These bucks have been genotyped since at least 1998, so they have very accurate genetic values derived from goats of the entire population, regardless of the breeding system. Access to a larger population of candidate bucks was not possible because adequate numbers of bucks were needed both in the reference population and as newborn candidate bucks at least 5 yr before the end of the performance data set in order

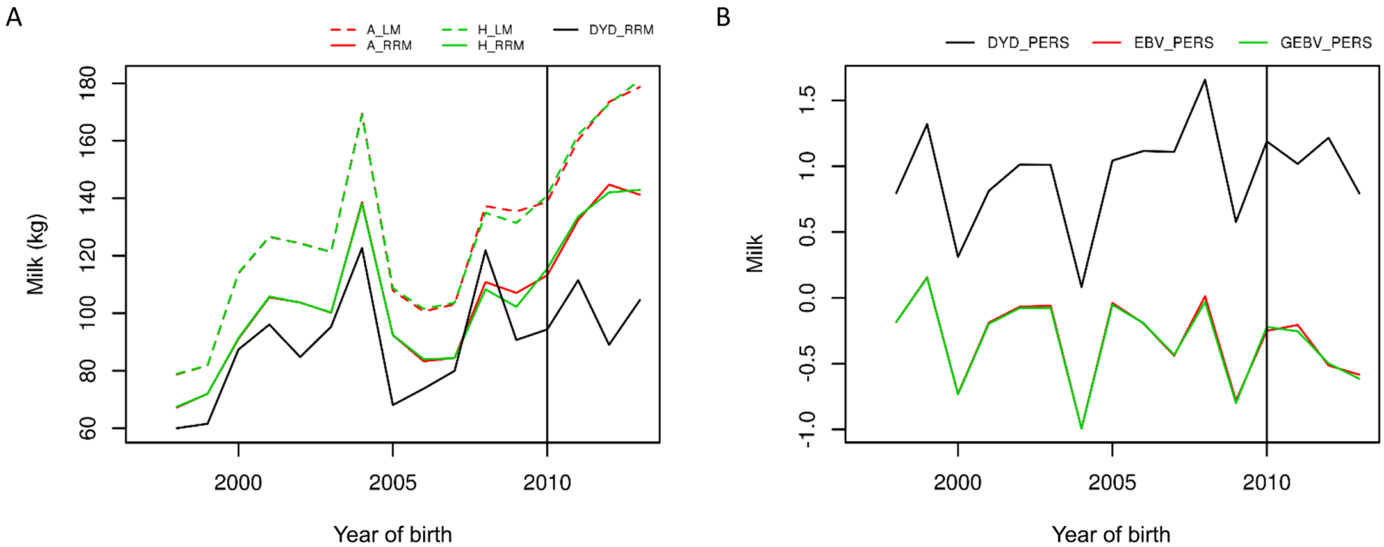

Figure 3. Plot of EBVs from the partial data set and daughter yield deviations (DYD) of genotyped bucks through years of birth (most of the time, the red curve is under the green curve). (A) Milk yield throughout the lactation EBV. (B) Milk persistency EBV. A_LM: lactation model without genomic information used to build the relationship matrix, H_LM: lactation model with genomic information used to build the relationship matrix, A_RRM: random regression model without genomic information used to build the relationship matrix, H_RRM: random regression model with genomic information used to build the relationship matrix, DYD_RRM: daughter yield deviations from random regression model, DYD_PERS: daughter yield deviations for persistency, EBV_PERS: estimated breeding value from RRM model without genomic information for persistency, GEBV_PERS: estimated breeding value obtained from the RRM with genomic information for persistency.

for these candidate males to have multiparous offspring. The analysis of the correlations between models showed that with the population considered, the gain from a LM_H to an RRM_H would be smaller than the gain from a LM_A to a LM_H. Another way to compare models is to study the correlations of buck rankings, as in Berry et al. (2011).

Analysis of SNPs' Effects

We chose to consider groups of 10 adjacent SNPs to study the percentage of variance explained because a smaller size of the windows may have led to smaller signals and bigger noise. and windows of more than 10 adjacent SNPs may combine certain QTLs. The genomic regions highlighted with LM on chromosome 6 and 19 were the same as those found in other French studies (Martin et al., 2017, 2018; Talouarn et al., 2020) on the same breed but using other methodologies (linkage association, linkage disequilibrium). The region around 26.1 Mb on chromosome 19 was also found by Mucha et al. (2018) for MY in crossbred goats (Saanen-Toggenburg-Alpine) who identified it as the ALOX12 gene linked to MY. The same region was also identified in a study from New Zealand in a mixed breed population composed by Saanen (Scholtens et al., 2020). This region is densely packed with genes, making it difficult to nominate candidate genes (Martin et al., 2018). The region on chromosome 6 is well known for casein (CSN1S1 and CSN2) and its association with protein yield and protein content (Martin et al., 2018). The LM and RRM for LEV point out the same regions. These results confirm that the 2 models are quite equivalent and led to similar percentages of variance. Differences between primiparous and multiparous animals were observed, but it is known that EBV are not the same traits in primiparous and in multiparous animals, as the genetic correlation between them is 0.69 (Arnal et al., 2020). Regions associated with persistency were the same as the regions associated with LEV in primiparous animals. Cardona et al. (2016) also reported in Creole goats that regions associated with caseins on chromosome 6 are associated with fat yield persistency too. In dairy cattle, Pryce et al. (2010), Strucken et al. (2012), and Kolbehdari et al. (2009) also found regions associated with persistency. They did not use persistency independent from LEV, as we did here, but instead used classes of DIM throughout the lactation. In dairy cattle, Macciotta et al. (2015) used the second eigenvector as we did to perform a GWAS but they did it at a phenotypic level. They found no region associated with the second eigenvector, possibly because they did not evaluate persistency at a genetic level. Studies in other dairy species have looked at the effect of genotypes on the shape of the lactation curve (Pauciullo et al., 2012; Szyda et al., 2014) based on different alleles, and found phenotypic differences in shape of the lactation curve according to genotypes. Here, the analysis of the SNPs'

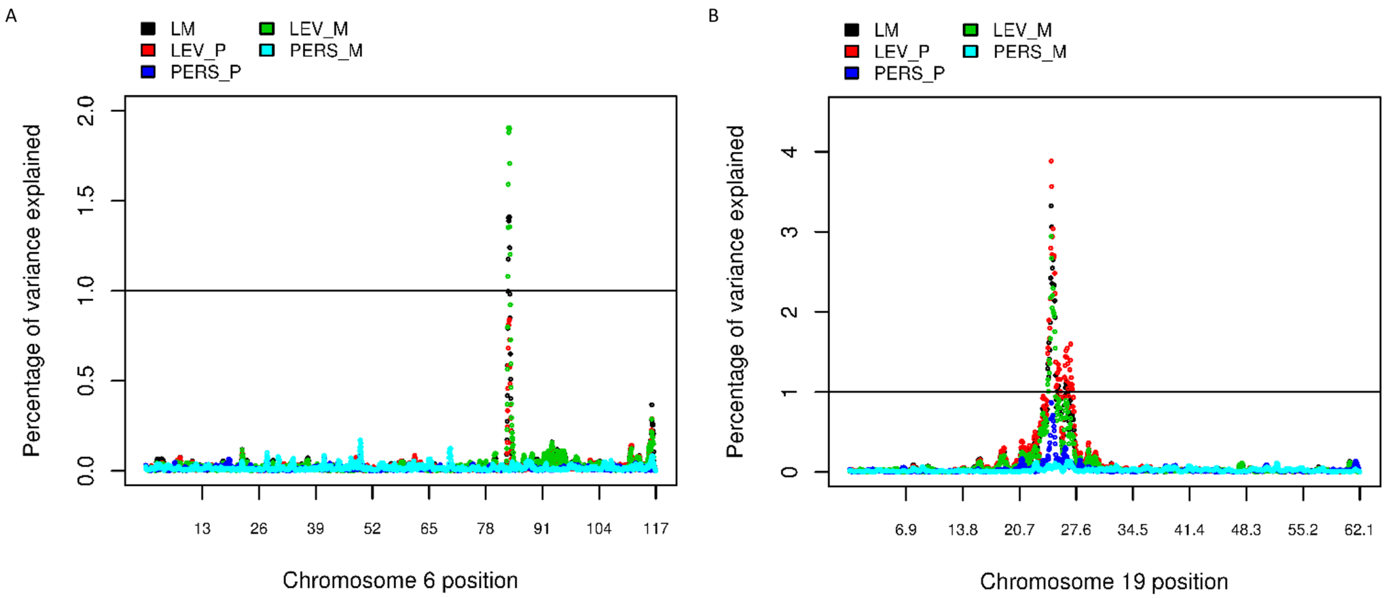

Figure 4. Percentages of total variance represented by segments of 10 adjacent SNPs for (A) chromosome 6 and (B) chromosome 19. LM: lactation model, LEV_P: first-parity production level, LEV_M: multiple-parity production level, PERS_P: first-parity persistency, PERS_M: multiple-parity persistency.

effects from ssGBLUP for LEV and PERS shows that some SNPs are more associated with LEV and PERS than others and that these SNPs would be useful to weigh in a genomic evaluation model. A ssWeightedG-BLUP RRM, as proposed in Karaman et al. (2018), could improve the accuracy of EBV on candidate bucks at birth.

CONCLUSIONS

An RRM using single-step genomic evaluation has been developed for dairy goats in France. The added genomic information increased correlations between DYD and EBV for candidate bucks. The LM and the RRM had fairly similar performances for correlations between candidate EBV and DYD. However, the ssG-BLUP RRM reduced the bias between DYD and EBV and improved the slope between DYD and EBV. This study shows that a single-step RRM is feasible to evaluate dairy goats in France and that it offers comparable performance to LM while adding a genomic evaluation of persistency, which is a trait of interest for dairy goat breeding. After the realization of this study, the French goat genetics team has arguments to change its LM for an RRM. Our analysis of SNPs' effects highlights that both LM and RRM found the same genomic regions associated with MY, including interesting genomic regions associated with persistency. Further investigations are required to confirm and refine these genomic regions.

ACKNOWLEDGMENTS

The first author received financial support from APIS-GENE (Paris, France) and the French National Association for Research and Technology (ANRT, Paris, France). This work was performed in the framework of the European Union's Horizon 2020 research and innovation program under Grant Agreement No 772787 (SMARTER). We thank Ignacy Misztal (University of Georgia, Athens, GA) and Andres Legarra (INRAE, UMR GenPhySE, Toulouse, France) for kindly opening access to blup90iod2 executable files, Marc Teissier (INRAE, UMR GenPhySE, Toulouse, France) for filtering the genomic data, and Andres Legarra (INRAE, UMR GenPhySE, Toulouse, France) for valuable and constructive discussions. The authors thank the anonymous reviewers for their valuable and constructive comments. The authors have not stated any conflicts of interest.

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752. https://doi.org/10.3168/jds.2009-2730.

Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. In Proceedings of the 10th World Congress of Genetics Applied to Livestock Production.

Arnal, M., H. Larroque, H. Leclerc, V. Ducrocq, and C. Robert-Granié. 2019. Genetic parameters for first lactation dairy traits in the Alpine and Saanen goat breeds using a random regression test-day model. Genet. Sel. Evol. 51:43. https://doi.org/10.1186/s12711-019-0485-3.

Arnal, M., H. Larroque, H. Leclerc, V. Ducrocq, and C. Robert-Granié. 2020. Estimation of genetic parameters for dairy traits and somatic cell score in the first 3 parities using a random regression test-day model in French Alpine goats. J. Dairy Sci. 103:4517–4531. https://doi.org/10.3168/jds.2019-17465.

Arnal, M., C. Robert-Franié, V. Ducrocq, and H. Larroque. 2023. JDS_ARNAL_SUPPLEMENTARY_MATERIAL_TABLE S1. figshare. Figure. https://doi.org/10.6084/m9.figshare.22592416.v3.

Baba, T., Y. Gotoh, S. Yamaguchi, S. Nakagawa, H. Abe, Y. Masuda, and T. Kawahara. 2017. Application of single-step genomic best linear unbiased prediction with a multiple-lactation random regression test-day model for Japanese Holsteins. Anim. Sci. J. 88:1226–1231. https://doi.org/10.1111/asj.12760.

Berry, D. P., R. D. Evans, and S. Mc Parland. 2011. Evaluation of bull fertility in dairy and beef cattle using cow field data. Theriogenology 75:172–181. https://doi.org/10.1016/j.theriogenology.2010.08.002.

Cardona, S. J. C., H. C. Cadavid, J. D. Corrales, S. Munilla, R. J. Cantet, and A. Rogberg-Muñoz. 2016. Longitudinal data analysis of polymorphisms in the κ-casein and β-lactoglobulin genes shows differential effects along the trajectory of the lactation curve in tropical dairy goats. J. Dairy Sci. 99:7299–7307. https://doi.org/10.3168/jds.2016-10954.

Clément, V., D. Boichard, A. Piacere, A. Barbat, and E. Manfredi. 2002. Genetic evaluation of French goats for dairy and type traits. In Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier.

Cole, J. B., and D. J. Null. 2009. Genetic evaluation of lactation persistency for five breeds of dairy cattle. J. Dairy Sci. 92:2248–2258. https://doi.org/10.3168/jds.2008-1825.

R2D2 Consortium, A. Fugeray-Scarbel, C. Bastien, M. Dupont-Nivet, and S. Lemarié. 2021. Why and how to switch to genomic selection: lessons from plant and animal breeding experience. Front. Genet. 12:629737. https://doi.org/10.3389/fgene.2021.629737.

Currò, S., C. L. Manuelian, M. De Marchi, P. De Palo, S. Claps, A. Maggiolino, G. Campanile, D. Rufrano, A. Fontana, G. Pedota, and G. Neglia. 2019. Autochthonous dairy goat breeds showed better milk quality than Saanen under the same environmental conditions. Arch. Anim. Breed. 62:83–89. https://doi.org/10.5194/aab-62-83-2019.

Druet, T., F. Jaffrézic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. J. Dairy Sci. 86:2480–2490. https://doi.org/10.3168/jds.S0022-0302(03)73842-9.

Druet, T., F. Jaffrézic, and V. Ducrocq. 2005. Estimation of genetic parameters for test day records of dairy traits in the first three lactations. Genet. Sel. Evol. 37:257. https://doi.org/10.1186/1297-9686-37-4-257.

Ducrocq, V., and M. P. Schneider. 2007. Generalization of the information source method to compute reliabilities in test models. Interbull Bull. 37:82.

Ducrocq, V. 1998. Genekit, BLUP software; version October 30, 2017. INRA SGQA.

Garcia-Baccino, C. A., A. Legarra, O. F. Christensen, I. Misztal, I. Pocrnic, Z. G. Vitezica, and R. J. Cantet. 2017. Metafounders are related to F_{st} fixation indices and reduce bias in single-step genomic evaluations. Genet. Sel. Evol. 49:34. https://doi.org/10.1186/s12711-017-0309-2.

ICAR. 2018. Guidelines for performance recording in dairy sheep and dairy goats. Accessed Apr. 12, 2023. https://www.icar.org/Guidelines/16-Dairy-Sheep-and-Goats.pdf.

ICAR. 2022. ICAR guidelines. Accessed Oct. 12, 2022. https://www.icar.org/index.php/icar-recording-guidelines/.

Kang, H., C. Ning, L. Zhou, S. Zhang, Q. Yan, and J.-F. Liu. 2018. Single-step genomic evaluation of milk production traits using multiple-trait random regression model in Chinese Holsteins. J. Dairy Sci. 101:11143–11149. https://doi.org/10.3168/jds.2018-15090.

Kang, H., L. Zhou, R. Mrode, Q. Zhang, and J. F. Liu. 2017. Incorporating the single-step strategy into a random regression model to enhance genomic prediction of longitudinal traits. Heredity 119:459–467. https://doi.org/10.1038/hdy.2016.91.

Karaman, E., M. S. Lund, M. T. Anche, L. Janss, and G. Su. 2018. Genomic prediction using multi-trait weighted GBLUP accounting for heterogeneous variances and covariances across the genome. G3 (Bethesda) 8:3549–3558.

Koivula, M., I. Strandén, J. Pösö, G. P. Aamand, and E. A. Mäntysaari. 2015. Single-step genomic evaluation using multitrait random regression model and test-day data. J. Dairy Sci. 98:2775–2784. https://doi.org/10.3168/jds.2014-8975.

Kolbehdari, D., Z. Wang, J. R. Grant, B. Murdoch, A. Prasad, Z. Xiu, E. Marques, P. Stothard, and S. S. Moore. 2009. A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. J. Anim. Breed. Genet. 126:216–227. https://doi.org/10.1111/j.1439-0388.2008.00793.x.

Larroque, H., J.-M. Astruc, A. Barbat, F. Barillet, D. Boichard, B. Bonaiti, V. Clément, I. David, G. Lagriffoul, I. Palhiere, A. Piacère, C. Robert-Granie, and R. Rupp. 2011. National Genetic Evaluations in Dairy Sheep and Goats in France. Wageningen Academic.

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., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. Genetics 200:455–468. https://doi.org/10.1534/genetics.115.177014.

Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. Genet. Sel. Evol. 50:53. https://doi.org/10.1186/s12711-018-0426-6.

Macciotta, N. P. P., G. Gaspa, L. Bomba, D. Vicario, C. Dimauro, M. Cellesi, and P. Ajmone-Marsan. 2015. Genome-wide association analysis in Italian Simmental cows for lactation curve traits using a low-density (7K) SNP panel. J. Dairy Sci. 98:8175–8185. https://doi.org/10.3168/jds.2015-9500.

Martin, P., I. Palhière, 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. 7:1872. https://doi.org/10.1038/s41598-017-02052-0.

Martin, P., I. Palhière, 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.

Misztal, I., D. Bradford, D. Lourenco, S. Tsuruta, Y. Masuda, A. Legarra, and T. J. Lawlor. 2017. Studies on inflation of GEBV in single-step GBLUP for type. Interbull Bull. 51:38–42.

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Pages 743–744 in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production.

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.

Oliveira, H. R., L. F. Brito, D. A. L. Lourenco, F. F. Silva, J. Jamrozik, L. R. Schaeffer, and F. S. Schenkel. 2019a. Invited review: Advances and applications of random regression models: From quantitative genetics to genomics. J. Dairy Sci. 102:7664–7683. https://doi.org/10.3168/jds.2019-16265.

Oliveira, H. R., L. F. Brito, F. F. Silva, D. A. L. Lourenco, J. Jamrozik, and F. S. Schenkel. 2019b. Genomic prediction of lactation curves for milk, fat, protein, and somatic cell score in Holstein cattle. J. Dairy Sci. 102:452–463. https://doi.org/10.3168/jds.2018-15159.

Oliveira, H. R., D. A. L. Lourenco, Y. Masuda, I. Misztal, S. Tsuruta, J. Jamrozik, L. F. Brito, F. F. Silva, J. P. Cant, and F. S. Schenkel. 2019c. Single-step genome-wide association for longitudinal traits of Canadian Ayrshire, Holstein, and Jersey dairy cattle. J. Dairy Sci. 102:9995–10011. https://doi.org/10.3168/jds.2019-16821.

Oliveira, H. R., D. A. L. Lourenco, Y. Masuda, I. Misztal, S. Tsuruta, J. Jamrozik, L. F. Brito, F. F. Silva, and F. S. Schenkel. 2019d. Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. J. Dairy Sci. 102:2365–2377. https://doi.org/10.3168/jds.2018-15466.

Pauciullo, A., G. Cosenza, R. Steri, A. Coletta, L. Jemma, M. Feligini, D. Di Berardino, N. P. Macciotta, and L. Ramunno. 2012. An association analysis between OXT genotype and milk yield and flow in Italian Mediterranean river buffalo. J. Dairy Res. 79:150–156. https://doi.org/10.1017/S0022029911000914.

Pryce, J. E., M. Haile-Mariam, K. Verbyla, P. J. Bowman, M. E. Goddard, and B. J. Hayes. 2010. Genetic markers for lactation persistency in primiparous Australian dairy cows. J. Dairy Sci. 93:2202–2214. https://doi.org/10.3168/jds.2009-2666.

Sargent, F. D., V. H. Lytton, and O. G. Wall Jr.. 1968. Test interval method of calculating dairy herd improvement association records. J. Dairy Sci. 51:170–179. https://doi.org/10.3168/jds.S0022-0302(68)86943-7.

Schaeffer, L. R., and J. Jamrozik. 2008. Random regression models: a longitudinal perspective. J. Anim. Breed. Genet. 125:145–146. https://doi.org/10.1111/j.1439-0388.2008.00748.x.

Schaeffer, L. R., J. Jamrozik, G. J. Kistemaker, and J. Van Doormaal. 2000. Experience with a test-day model. J. Dairy Sci. 83:1135–1144. https://doi.org/10.3168/jds.S0022-0302(00)74979-4.

Scholtens, M., A. Jiang, A. Smith, M. Littlejohn, K. Lehnert, R. Snell, N. Lopez-Villalobos, D. Garrick, and H. Blair. 2020. Genome-wide association studies of lactation yields of milk, fat, protein and somatic cell score in New Zealand dairy goats. J. Anim. Sci. Biotechnol. 11:55. https://doi.org/10.1186/s40104-020-00453-2.

Strucken, E. M., R. H. Bortfeldt, D. J. De Koning, and G. A. Brockmann. 2012. Genome-wide associations for investigating time-dependent genetic effects for milk production traits in dairy cattle. Anim. Genet. 43:375–382. https://doi.org/10.1111/j.1365-2052.2011.02278.x.

Szyda, J., J. Komisarek, and I. Antkowiak. 2014. Modelling effects of candidate genes on complex traits as variables over time. Anim. Genet. 45:322–328. https://doi.org/10.1111/age.12144.

Talouarn, E., P. Bardou, I. Palhière, C. Oget, V. Clément, G. Tosser-Klopp, R. Rupp, and C. Robert-Granié. 2020. Genome wide association analysis on semen volume and milk yield using different strategies of imputation to whole genome sequence in French dairy goats. BMC Genet. 21:19. https://doi.org/10.1186/s12863-020-0826-9.

Täubert, H., Z. Liu, J. Tarrès, and V. Ducrocq. 2010. An approach to compute EDC and DYD for test-day models. Proceedings of the 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany.

Teissier, M., H. Larroque, and C. Robert-Granie. 2019. Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in French dairy goats. J. Dairy Sci. 102:3142–3154. https://doi.org/10.3168/jds.2018-15650.

Tosser-Klopp, G., P. Bardou, O. Bouchez, C. Cabau, R. Crooijmans, Y. Dong, C. Donnadieu-Tonon, A. Eggen, H. C. 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, W. Jun Wang, Wang, and W. Zhang. 2014. Design and characterization of a 52K SNP chip for goats. PLoS One 9:e86227. https://doi.org/10.1371/journal.pone.0086227.

VanRaden, P. M., and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. J. Dairy Sci. 74:2737–2746. https://doi.org/10.3168/jds.S0022-0302(91)78453-1.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, R. L. Fernando, Z. Vitezica, R. Okimoto, T. Wing, R. Hawken, and W. M. Muir. 2014. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front. Genet. 5:134. https://doi.org/10.3389/fgene.2014.00134.

ORCIDS

M. Arnal https://orcid.org/0000-0002-2671-5026
C. Robert-Granié https://orcid.org/0000-0001-5313-2187
V. Ducrocq https://orcid.org/0000-0002-1503-5199