

## Research Article

\*Present address: Council on Dairy Cattle Breeding – CDCB, Bowie, Maryland, USA.

**Cite this article:** Paiva JT, Mota RR, Lopes PS, Hammami H, Vanderick S, Oliveira HR, Veroneze R, Silva FF and Gengler N (2022). Genomic prediction and genetic correlations estimated for milk production and fatty acid traits in Walloon Holstein cattle using random regression models. *Journal of Dairy Research* **89**, 222–230. <https://doi.org/10.1017/S0022029922000474>

Received: 31 October 2021

Revised: 5 April 2022

Accepted: 26 April 2022

First published online: 5 September 2022

**Keywords:**

Genetic correlation; genomic prediction; MIR; single-step GBLUP; test-day model

**Author for correspondence:**

José Teodoro Paiva,

Email: [teo.paiva@hotmail.com](mailto:teo.paiva@hotmail.com)

# Genomic prediction and genetic correlations estimated for milk production and fatty acid traits in Walloon Holstein cattle using random regression models

José Teodoro Paiva<sup>1</sup>, Rodrigo Reis Mota<sup>2,\*</sup>, Paulo Sávio Lopes<sup>1</sup>, Hedi Hammami<sup>2</sup>, Sylvie Vanderick<sup>2</sup>, Hinayah Rojas Oliveira<sup>3</sup>, Renata Veroneze<sup>1</sup>, Fabyano Fonseca e Silva<sup>1</sup> and Nicolas Gengler<sup>2</sup>

<sup>1</sup>Department of Animal Sciences, Universidade Federal de Viçosa, Viçosa, MG, Brazil; <sup>2</sup>Gembloux Agro-Bio Tech, University of Liège, TERRA Teaching and Research Centre, B-5030 Gembloux, Belgium and <sup>3</sup>Department of Animal Biosciences, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, Canada

**Abstract**

The aims of this study were to: (1) estimate genetic correlation for milk production traits (milk, fat and protein yields and fat and protein contents) and fatty acids (FA: C16:0, C18:1 *cis*-9, LCFA, SFA, and UFA) over days in milk, (2) investigate the performance of genomic predictions using single-step GBLUP (ssGBLUP) based on random regression models (RRM), and (3) identify the optimal scaling and weighting factors to be used in the construction of the **H** matrix. A total of 302 684 test-day records of 63.875 first lactation Walloon Holstein cows were used. Positive genetic correlations were found between milk yield and fat and protein yield ( $r_g$  from 0.46 to 0.85) and between fat yield and milk FA ( $r_g$  from 0.17 to 0.47). On the other hand, negative correlations were estimated between fat and protein contents ( $r_g$  from  $-0.22$  to  $-0.59$ ), between milk yield and milk FA ( $r_g$  from  $-0.22$  to  $-0.62$ ), and between protein yield and milk FA ( $r_g$  from  $-0.11$  to  $-0.19$ ). The selection for high fat content increases milk FA throughout lactation ( $r_g$  from 0.61 to 0.98). The test-day ssGBLUP approach showed considerably higher prediction reliability than the parent average for all milk production and FA traits, even when no scaling and weighting factors were used in the **H** matrix. The highest validation reliabilities ( $r^2$  from 0.09 to 0.38) and less biased predictions ( $b_1$  from 0.76 to 0.92) were obtained using the optimal parameters (i.e.,  $\omega = 0.7$  and  $\alpha = 0.6$ ) for the genomic evaluation of milk production traits. For milk FA, the optimal parameters were  $\omega = 0.6$  and  $\alpha = 0.6$ . However, biased predictions were still observed ( $b_1$  from 0.32 to 0.81). The findings suggest that using ssGBLUP based on RRM is feasible for the genomic prediction of daily milk production and FA traits in Walloon Holstein dairy cattle.

The triacylglycerols (TAG) account for approximately 95% of the lipid fraction in milk solids and are composed of fatty acids (FAs) of different lengths and saturations (Haug *et al.*, 2007). Interest in milk FA profile is increasing given its important nutritional value for human health and its influence on the technological properties of milk and dairy products, such as the spreadability of butter. In addition, increased levels of unsaturated FA (UFA) content in milk can contribute to lower stability and oxidation, which directly impact dairy products (Hanus *et al.*, 2018). On the other hand, C18:1 *cis*-9 (oleic acid) is the unsaturated FA with the highest concentration in milk, which is considered to be favorable for human health (Haug *et al.*, 2007). Notwithstanding, Jorjong *et al.* (2014) reported that C18:1 *cis*-9 shows potential as an early warning biomarker for metabolic status and subclinical ketosis in dairy cows. Studies have shown that the milk FA profile is under genetic control in dairy cattle (Bastin *et al.*, 2011; Paiva *et al.*, 2022). Thus, changes in FA composition can be made through genetic selection. Moreover, the availability of these specialized phenotypes makes their inclusion in genetic and genomic evaluations possible. This brings unprecedented and substantial impacts to improving milk quality.

However, potential implications of selection for milk production and FA traits require accurate estimation of genetic parameters in the studied population. Evaluating the impact of selecting one trait instead of another is essential to predict indirect genetic gains to determine the best selection strategies in a breeding program. In this context, to the best of our knowledge, there are few studies evaluating the impact of the selection of milk production traits on the FA profile over days in milk (e.g., Bastin *et al.*, 2011; Fleming *et al.*, 2018). It is important to mention that previous studies did not perform genomic analysis and instead were only pedigree-based. The inclusion of genomic information has allowed an improvement

of genomic estimated breeding value (GEBV) accuracies over the traditional pedigree-based estimated breeding values (EBV; Legarra *et al.*, 2014). Single-step genomic best linear unbiased prediction (ssGBLUP) is a unified approach that optimally combines phenotypic records, pedigree information, and genomic information in the calculation of GEBV (Aguilar *et al.*, 2010; Christensen and Lund, 2010). Currently, given the limitation of genotyped animals in real populations, ssGBLUP is among the most efficient methods and has been routinely used in animal breeding (Lourenco *et al.*, 2020). However, this approach usually requires the use of optimal scaling factors (Misztal *et al.*, 2017; Lourenco *et al.*, 2020) to assure the ideal compatibility between the marker-based relationship and pedigree-based relationship matrices. Adjusting the genomic relationship matrix toward their expected values in the pedigree-based matrix can improve the accuracy and reduce bias of GEBV.

In general, studies using ssGBLUP based on RRM provided higher accuracy and less biased GEBV compared with other methods (Koivula *et al.*, 2015; Oliveira *et al.*, 2019). To the best of our knowledge, no study has investigated the optimal scaling factors to perform ssGBLUP evaluation for milk FAs using RRM in dairy cattle. Therefore, the aims of this study were to: (1) estimate genetic correlation for milk production and FA traits over days in milk using pedigree and genomic information; (2) investigate the performance of genomic prediction (in terms of reliability and bias) using the ssGBLUP approach and compare it with the pedigree-based method in Walloon Holstein cattle; and (3) identify the optimal scaling factors to be used in the construction of the **H** matrix in this population.

## Material and methods

### Datasets

Phenotypic and genotypic data were extracted from the Walloon genetic evaluation performed in Belgium. Milk samples (containing 50% morning milk and 50% evening milk) were routinely collected by the Walloon Breeders Association (AWE; Ciney, Belgium) and were analyzed by a mid-infrared MilkoScan FT6000 spectrometer (Foss, Hillerød, Denmark). Fat yield (kg), protein yield (kg), fat content (%), protein content (%), C16:0 fatty acid (palmitic acid), C18:1 *cis*-9 fatty acid (oleic acid), long-chain fatty acid (LCFA), saturated fatty acids (SFA), and group unsaturated fatty acids (UFA) were the traits predicted from the test-day milk samples evaluated in this study in addition to daily milk yield (kg). More details about the Walloon Holstein cow population and dataset quality control can be found in Paiva *et al.* (2022). The final dataset is comprised of 302 684 test-day records from 63 875 first-parity Walloon Holstein cows, raised in 856 herds. Descriptive statistics of milk production and FA traits are shown in the Supplementary Table S1.

Routine evaluation genotypes for Walloon Holsteins came from 11 different SNP chips (Illu50K, Illu50K2, Illu50K3, IlluHD, GGP150K, IDB3, EuroG10K, IlluHD3, EuroG\_MD) of different densities. Therefore, imputation and edits were performed to develop a list of single nucleotide polymorphism (SNP) markers. The current list contains 19 468 SNPs in common with all 11 chips in the database. In this study, a total of 5057 animals were genotyped (620 sires and 1931 dams in the pedigree). Genotypic quality control was performed using the preGSf90 software (Aguilar *et al.*, 2014). SNP markers with a minor allele frequency lower than 0.05 ( $n = 1,167$ ), call rate lower than 0.90

( $n = 494$ ), monomorphic ( $n = 0$ ), and SNP with Mendelian conflicts ( $n = 3$ ) were removed. Likewise, individual samples with call rates lower than 0.90 ( $n = 0$ ) and with parent-progeny Mendelian conflicts ( $n = 20$ ) were also excluded. After quality control, the genotypic dataset included 5037 genotyped animals (2865 cows with test-day records) and 17 866 SNPs, which were used for further analysis.

### Statistical model

The random regression test-day bi-trait model (RRM) was based on the optimal statistical model found in a previous study by Paiva *et al.* (2022):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Q}(\mathbf{W}\mathbf{h} + \mathbf{Z}\mathbf{a} + \mathbf{Z}\mathbf{p}) + \mathbf{e}, \quad (1)$$

where  $\mathbf{y}$  is the vector of phenotypic records for each trait;  $\boldsymbol{\beta}$  is the vector of systematic effects, which included herd  $\times$  test-day, gestation stage, minor lactation stage, and major lactation stage  $\times$  age at calving  $\times$  season of calving;  $\mathbf{h}$  is the vector of herd-year of calving random regression coefficients;  $\mathbf{a}$  is the vector of additive genetic random regression coefficients;  $\mathbf{p}$  is the vector of permanent environment random regression coefficients;  $\mathbf{X}$ ,  $\mathbf{W}$ , and  $\mathbf{Z}$  are the incidence matrices assigning the observations to effects;  $\mathbf{Q}$  is the covariate matrix for the third-order Legendre polynomials; and  $\mathbf{e}$  is the vector of random residuals.

The expectations and covariance structure for the random effects were given by:

$$E[\mathbf{y}] = \begin{bmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \text{ and } \text{Var} \begin{bmatrix} \mathbf{h} \\ \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{H}\mathbf{Y}_0 \otimes \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_0 \otimes \mathbf{H} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_0 \otimes \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_0 \otimes \mathbf{I} \end{bmatrix} \quad (2)$$

where  $\mathbf{H}\mathbf{Y}_0$ ,  $\mathbf{G}_0$ , and  $\mathbf{P}_0$  are the random regression coefficients (co) variances matrices for herd-year of calving, additive genetic, and permanent environment effects, respectively;  $\mathbf{I}$  is the identity matrix;  $\mathbf{H}$  is a matrix that combines pedigree and genomic information;  $\mathbf{R}_0$  is the (co)variance matrix of residual effects among traits, assumed to be homogeneous over the lactation; and  $\otimes$  is the Kronecker product.

The inverse of the **H** matrix was created as Aguilar *et al.* (2010), and Christensen and Lund (2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau(\alpha\mathbf{G} + \beta\mathbf{A}_{22})^{-1} - \omega\mathbf{A}_{22}^{-1} \end{bmatrix}, \quad (3)$$

where  $\mathbf{A}^{-1}$  is the inverse of the pedigree-based numerator relationship matrix,  $\mathbf{G}$  is the genomic relationship matrix, and  $\mathbf{A}_{22}^{-1}$  is the proportion of  $\mathbf{A}^{-1}$  related to the genotyped animals. The genomic relationship matrix was created as proposed by VanRaden (2008):

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{2 \sum p_i(1-p_i)}, \quad (4)$$

where  $\mathbf{Z} = \mathbf{M} - \mathbf{P}$ ;  $\mathbf{M}$  is a matrix of genotypes centered for allele frequencies (i.e., 0, 1 and 2 to represent *aa*, *Aa* and *AA*,

respectively), with dimensions equal to a number of animals by number of SNP;  $\mathbf{P}$  contains the allele frequency for SNP  $i$  ( $p_i$ ) in its column  $i$ th, expressed as  $2(p_i - 0.5)$ , and  $2 \sum p_i(1-p_i)$  is a scaling parameter.

### Genetic correlation estimation

The genetic correlations were estimated through bi-trait analysis *via* the Gibbs sampling algorithm using GIBBS2F90 software (Misztal *et al.*, 2002). A chain length of 250 000 iterations and convergence criteria were used similar to Paiva *et al.* (2022). The genetic (co)variance matrix for all DIM was obtained according to Druet *et al.* (2003), which is described as:

$$\mathbf{G}^* = \mathbf{Q}\mathbf{G}_0\mathbf{Q}', \quad (5)$$

where  $\mathbf{G}^*$  is a genetic (co)variance matrix among the traits for all DIM ranging from 5 to 305 d;  $\mathbf{G}_0$  is a (co)variance matrix of genetic regression coefficients; and  $\mathbf{Q}$  is a matrix with the values of the three coefficients of the third-order Legendre polynomial for each DIM.

The posterior marginal distribution samples for genetic correlation ( $r_{gj}$ ) at test-day  $j$  were calculated as follows:

$$r_{gj} = \frac{\sigma_{a12j}}{\sqrt{\sigma_{a1j}^2 \times \sigma_{a2j}^2}}, \quad (6)$$

where  $\sigma_{a12j}$  is the additive genetic covariance between traits 1 and 2 at test-day  $j$ , and  $\sigma_{a1j}^2$  and  $\sigma_{a2j}^2$  are the additive genetic variances of traits 1 and 2, respectively, at test-day  $j$ .

### Breeding value prediction

Breeding values for random regression coefficients of each animal were predicted using both the traditional BLUP ( $\mathbf{A}$  matrix) and the ssGBLUP ( $\mathbf{H}$  matrix) approaches on single-trait analysis. The optimal values for the scaling factors ( $\tau=1$  and  $\omega=0.70$ ) reported by Tsuruta *et al.* (2011) and used by Colinet *et al.* (2017), in a subset of the same Walloon Holstein dairy population, were used in this study for the genomic evaluation of milk production traits. In summary, the  $\omega$  parameter helps to reduce GEBV inflation and compensates for pedigree incompleteness (Masuda *et al.*, 2016; Lourenco *et al.*, 2020). Moreover,  $\tau$  and  $\omega$  are used to account for the reduced genetic variance and for different depths of pedigree to make  $\mathbf{G}^{-1}$  compatible with  $\mathbf{A}_{22}^{-1}$  and  $\mathbf{A}^{-1}$ . The weighting factors ( $\alpha=0.60$  and  $\beta=0.40$ ) were used for milk production traits as suggested by Colinet *et al.* (2017) because these values reflect the partition of the genetic variance between the SNP markers and residual polygenic parts. The scaling and weighting factors described previously have been used in the official Walloon evaluation for milk production traits.

As changes in scaling and weighting factors were not investigated for milk FAs in the population analyzed in this study, scenarios with different values of scaling factors were tested for  $\omega$  (i.e., 0.6, 0.7, 0.8, 0.9, and 1.0). In this analysis, the proportion of markers variance was fixed to  $\alpha=0.95$ , and the polygenic variance was fixed to  $\beta=0.05$  (default values). Additionally, various scenarios with fixed  $\omega=1$  (default values) and different values of  $\alpha$  (i.e., 0.6, 0.7, 0.8, 0.95) were tested. It is worth noting that for all scenarios,  $\tau=1$  (default values) was used, since changes in this parameter have been reported to have small effects on the

bias of GEBV (Oliveira *et al.*, 2019). The best combination of these parameters for each trait analyzed was chosen according to the validation reliabilities and regression coefficients (see details in the 'Prediction Reliability and Bias' section below).

For each trait, EBV and GEBV of animal  $i$  at test-day  $j$  were obtained from a posteriori distribution of additive genetic (estimated by BLUP) and genomic (estimated by ssGBLUP) random regression coefficients as follows:

$$\text{EBV}_{ij} = \mathbf{C}_j\hat{\mathbf{a}}_i \text{ and } \text{GEBV}_{ij} = \mathbf{C}_j\hat{\mathbf{g}}_i \quad (7)$$

where  $\mathbf{C}_j$  is a matrix of independent covariates associated with the Legendre polynomials for test-day  $j$ ; and  $\hat{\mathbf{a}}_i$  and  $\hat{\mathbf{g}}_i$  are the vectors of predicted additive genetic and genomic coefficients for each animal  $i$ , respectively. The BLUPF90 software (Misztal *et al.*, 2002) was used to obtain the solutions of the additive genetic and genomic for the random regression coefficients.

### Prediction reliability and bias

A reduced dataset was created in which the last four years of the full data were cut off (i.e., excluding records of daughters of young bulls). Animals born until 2015 (285 507 test-day records from 60 292 cows, in which 2695 are genotyped) and between 2016 and 2019 (17 177 test-day records from 3,583 cows, in which 170 were genotyped) were grouped as the training and validation populations, respectively. Therefore, the phenotypes of animals born after 2015 were excluded from the analysis, to create a reduced dataset to be used to predict the GEBV using ssGBLUP and parent average (PA) using the traditional BLUP. This scenario allows us to mimic the Interbull GEBV test (Mäntysaari *et al.*, 2010). The full dataset was used to predict current EBV, which was used as a benchmark to validate GEBV and PA obtained from the reduced data set for validation animals by assessing the reliability and bias of genomic predictions. Genotyped bulls that only had daughters born between 2016 and 2019 were defined as validation animals (151 bulls). These 151 validation bulls had 3,583 daughters with a total of 17 177 test-day records. The total number of phenotyped and genotyped animals and test-day records in the full, training, and validation datasets are shown in Supplementary Table S2.

The validation reliability for each trait was calculated as the squared Pearson correlation coefficient ( $r^2$ ) between the daily GEBV estimated based on the reduced dataset (i.e., excluding phenotypes for the validation animals or their descendants) and the daily EBV estimated based on the full dataset, considering only animals in the validation population. The full dataset EBV has also been used to validate the performance of genomic predictions using RRM in other studies (Oliveira *et al.*, 2019; Freitas *et al.*, 2020). Likewise, only the bulls from the validation population were considered to assess the genomic prediction bias, which was calculated by obtaining the regression coefficient ( $b_1$ ) estimated using a linear regression of the daily EBV from the full dataset on GEBV from the reduced dataset (i.e.,  $\text{EBV} = b_0 + b_1 \times \text{GEBV}$ ) (Mäntysaari *et al.*, 2010). To evaluate the impact of including genomic information and to compare the prediction reliability and bias of GEBV to those of EBV from traditional genetic evaluation, the parent average (PA) was predicted for validation animals. The PA was used to calculate  $r^2$  (i.e., squared Pearson correlation coefficient between EBV and PA) and  $b_1$  (i.e., obtained from  $\text{EBV} = b_0 + b_1 \times \text{PA}$ ) using daily PA and daily EBV predicted based on the reduced and full dataset, respectively.

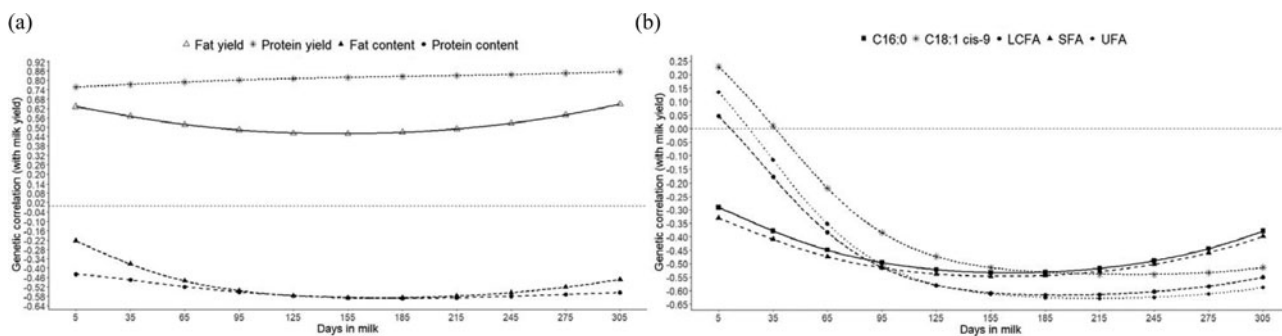


Fig. 1. Posterior means of genetic correlations over days in milk estimated between milk yield and (a) milk production traits and (b) FA traits.

Results

Genetic correlation estimates

Figure 1a shows the genetic correlations of fat and protein yield with milk yield (range from 0.46 to 0.85) over days in milk, as well as of fat and protein contents (range from -0.22 to -0.59). Posterior mean genetic correlations of FA with milk yield over days in milk are shown in Figure 1b. In general, most estimates were negative, and higher magnitudes were observed in the middle of lactation. The genetic correlation curves of milk yield with fat yield and content, C16:0, and SFA had similar patterns across lactation, with estimates ranging from -0.22 to -0.59.

Positive genetic correlations were estimated between fat yield and milk FA, and they tended to decrease across lactation, as shown in Figure 2a. Higher estimates were found for C16:0 (range from 0.43 to 0.49), but for other FAs the estimate curves were similar, with the lowest values at 305 DIM. Genetic correlations of milk FA with protein yield presented a similar pattern over days in milk, as shown in Figure 2b. Positive estimates were found at the beginning of lactation (35-65 DIM) and

became negative thereafter, with average daily genetic correlations ranging from -0.11 to -0.19.

The posterior mean genetic correlations estimated between milk FA and fat content were positive across lactation and had a similar pattern, as shown in Figure 2c. The estimates increased until the middle of lactation where maximum values were reached. Then, they were stable until the end of lactation. Genetic correlations estimated between fat content and C16:0 as well as SFA were similar and high, ranging from 0.89 to 0.98 across lactation. On the other hand, lower estimates were found between fat content and protein content (range from 0.55 to 0.72) and C18:1 cis-9 (range from 0.44 to 0.66).

Genetic correlations estimated between protein content and milk FA are shown in Figure 2d. Estimates were positive throughout lactation (range from 0.06 to 0.74) and showed the same trend of low values in early lactation, increasing thereafter. From the middle to the end of lactation, estimated genetic correlations between protein content and C16:0, LCFA, SFA, and UFA were moderate to high and slightly close over days in milk, ranging from 0.56 to 0.74.

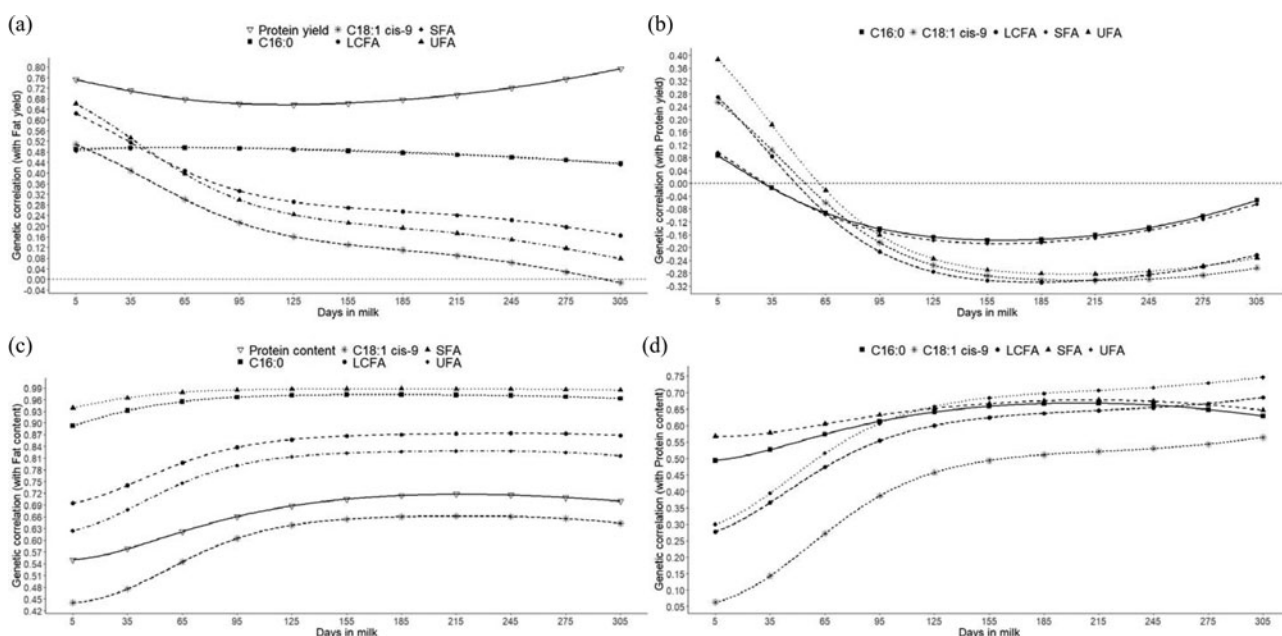


Fig. 2. Posterior means of genetic correlations over days in milk estimated between milk production and FA traits and (a) fat yield, (b) protein yield, (c) fat content, and (d) protein content.



Genetic correlations among all FA traits were positive over days in milk, with higher magnitude estimates observed from the middle (around 125 DIM) to the end of lactation, as shown in Figure 3. The highest genetic correlations for FA were found between C16:0 and SFA, C18:1 *cis*-9 and LCFA, C18:1 *cis*-9 and UFA, and between LCFA and UFA, which tended to be slightly stable along the lactation curve. Genetic correlations among the other FAs had similar patterns, with low values in early lactation that increased as lactation progressed up to the middle of lactation. Genetic correlations estimated between C18:1 *cis*-9 and C16:0 and between C18:1 *cis*-9 and SFA were similar, ranging from 0.07 to 0.53. Moreover, similar genetic correlations across lactation were found between C16:0 and LCFA (ranging from 0.35 to 0.76) and between C16:0 and UFA (ranging from 0.25 to 0.70). Likewise, the genetic correlation estimated between SFA and LCFA and between SFA and UFA ranged from 0.41 to 0.78 and from 0.31 to 0.72, respectively.

### Prediction reliability and bias

The average validation reliability and bias (with their respective standard deviations) of EBV and GEBV estimated for milk production and FA traits are shown in Tables 1 and 2, respectively. Overall, the use of ssGBLUP increased the reliability compared to the traditional BLUP (i.e., parent average) for the validation bulls even when no scaling and weighting factors (default;  $\omega = 1$

and  $\alpha = 0.95$ ) were used for milk production traits. However, the highest validation reliabilities were obtained using scaling and weighting factors  $\omega = 0.7$  and  $\alpha = 0.6$  (Table 1) to combine  $\mathbf{G}^{-1}$  and  $\mathbf{A}_{22}^{-1}$  in the genomic evaluations for milk production traits. Average reliabilities were especially higher for the milk yield, fat and protein contents ( $r^2$ ; 0.38, 0.19, and 0.18, respectively), while for fat and protein yield, they tended to be lower ( $r^2$ ; 0.14 and 0.09, respectively). In addition, the use of optimal scaling factors yielded the least biased prediction for milk production traits, with average  $b_1$  coefficients ranging from 0.76 to 0.92 (Table 1).

For milk FAs, the ssGBLUP models achieved higher average validation reliability than the parent average (Table 2). Similar validation reliabilities were estimated using different scaling and weighting factors for all milk FAs. The largest gain in genomic reliability was observed for C16:0 and SFA ( $r^2$ ; 0.11 and 0.17, respectively), as well as in all the scenarios considered. The validation average reliability was low for C18:1 *cis*-9 (0.04), LCFA (0.05) and UFA (0.05). However, a greater effect of scaling and weighting factors was observed in the regression coefficients for milk FA (Table 2). The inclusion of optimal values of  $\omega = 0.6$  and  $\alpha = 0.6$  for genomic evaluation of milk FAs greatly decreased bias (i.e., average  $b_1$  coefficients ranged from 0.32 to 0.81). Even though an improvement was observed, the GEBV predicted for most milk FA traits was still inflated when using the optimal scaling and weighting factors.

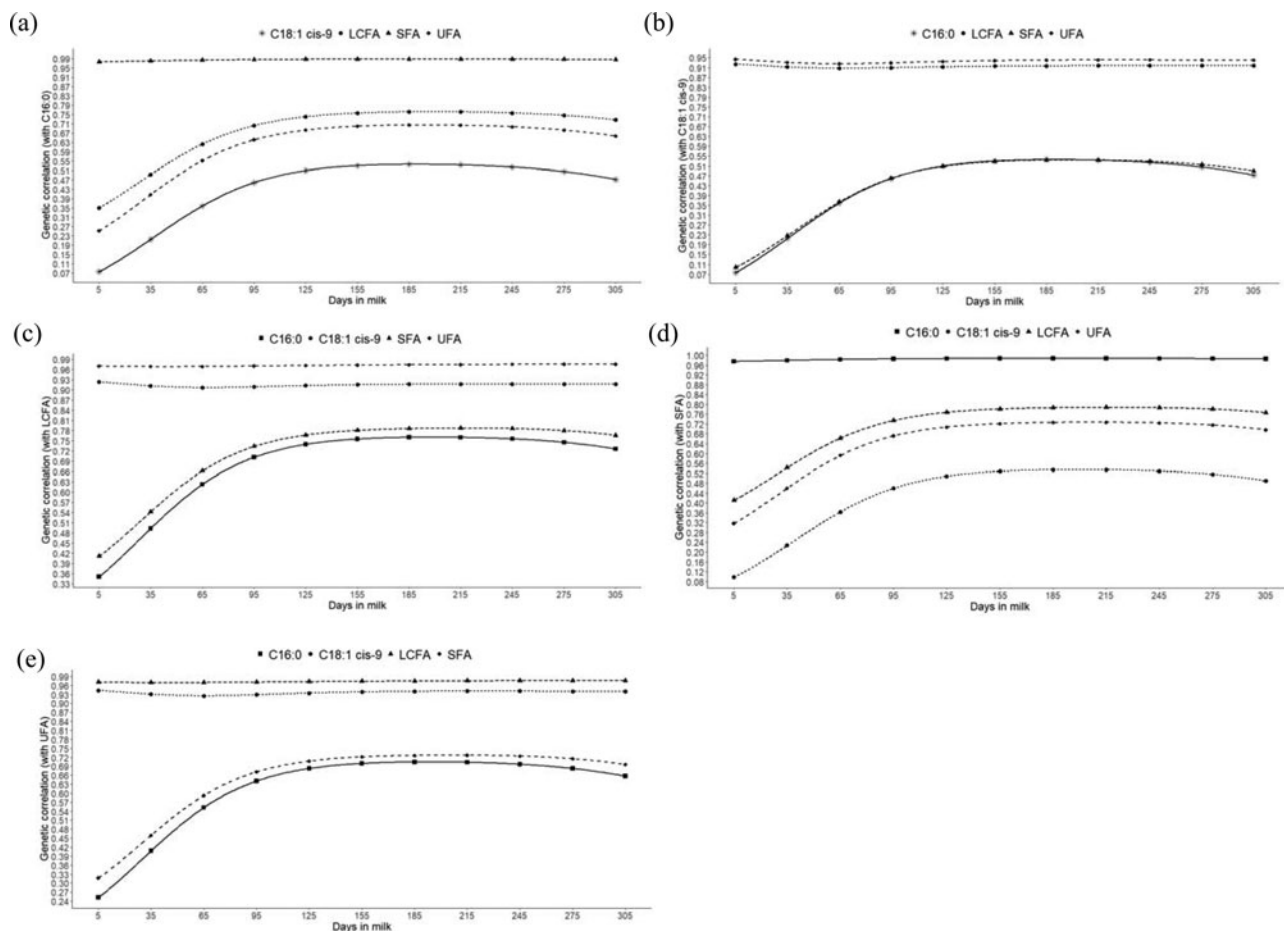


Fig. 3. Posterior means of genetic correlations over days in milk among milk fatty acid traits with (a) C16:0, (b) C18:1 *cis*-9, (c) LCFA, (d) SFA, and (e) UFA.

**Table 1.** Mean validation reliabilities ( $r^2$ ), regression coefficients ( $\hat{b}_1$ ) and their respective standard deviation for parentage average (PA) and genomic breeding value estimated assuming different\* scaling ( $\omega$ ) and weighting ( $\alpha$ ) factors for milk production traits for the Walloon Holstein bulls.

Scenario	Milk yield		Fat yield		Protein yield		Fat content		Protein content	
	$r^2$	$\hat{b}_1$	$r^2$	$\hat{b}_1$	$r^2$	$\hat{b}_1$	$r^2$	$\hat{b}_1$	$r^2$	$\hat{b}_1$
PA	0.35 (0.04)	0.91 (0.15)	0.12 (0.03)	0.70 (0.13)	0.07 (0.03)	0.71 (0.15)	0.12 (0.02)	0.71 (0.08)	0.10 (0.04)	0.73 (0.19)
$\omega$ 1 $\alpha$ 0.95 <sup>a</sup>	0.36 (0.06)	0.91 (0.12)	0.14 (0.03)	0.47 (0.10)	0.08 (0.03)	0.59 (0.11)	0.16 (0.09)	0.65 (0.21)	0.16 (0.09)	0.60 (0.18)
$\omega$ 0.7 $\alpha$ 0.6 <sup>b</sup>	0.38 (0.04)	0.92 (0.15)	0.14 (0.04)	0.78 (0.16)	0.09 (0.04)	0.77 (0.15)	0.19 (0.09)	0.86 (0.38)	0.18 (0.10)	0.76 (0.24)

<sup>a</sup> $\omega = 1$  and  $\alpha = 0.95$ ; default values.

<sup>b</sup> $\omega = 0.70$  and  $\alpha = 0.60$ ; the optimal values assumed for scaling and weighting factors proposed by Colinet *et al.* (2017).

\* $\tau$  scaling factor used for genomic relationship matrix ( $1.0G^{-1}$ );  $\omega$  = scaling factor used for pedigree relationship matrix ( $\omega A_{22}^{-1}$ );  $\alpha$  and  $\beta$  = weighting factors used for polygenic effect (ssGBLUP -  $\beta A_{22}$ ).

## Discussion

### Genetic correlations

One century of selection has mostly focused on increasing milk production (Miglior *et al.*, 2017). However, an antagonistic association across lactation ( $r_g$  from  $-0.22$  to  $-0.59$ ) was found between milk yield and fat and protein content. Thus, herds selected to increase milk volume can disfavor simultaneous milk content genetic gain, although they will increase fat and protein yield. Hence, challenges arise and are likely to be overcome by cautious evaluation of selection criteria chosen in the dairy breeding program. Decisions should be made to achieve the main selection objectives, particularly based on the payment system for the farmer, recognizing that they could change according to each country.

In general, daily genetic correlations between milk yield and FA were negative (they ranged from  $-0.38$  to  $-0.49$ ), which indicates that selection for improved milk yield would affect to a greater or lesser extent all FA traits in milk throughout lactation. Although in early lactation the genetic correlations were positive and weak for LCFA, C18:1 *cis*-9 and UFA, the selection for higher milk yield would decrease FA in milk as soon as lactation progressed, thus showing a strong influence mostly on C18:1 *cis*-9, for which the correlation becomes negative after 35 DIM. As already reported by Bastin *et al.* (2011), a variation in the milk FA profile is prominent in early lactation, and genetic correlations with milk yield are weaker than in late lactation.

Highly positive genetic correlations between SFA and fat content (average above 0.96) and protein content (average above 0.62) were also reported by Penasa *et al.* (2015), who found genetic correlations of 0.99 and 0.60, respectively, in Holstein cows. In addition, C16:0 had a strong genetic correlation with fat content, which could be explained because of its important role in the synthesis of triacylglycerol in the mammary gland. Bovine milk fat is mainly composed of triacylglycerols (95%), and the most abundant fatty acid is C16:0. The increase in fat content has led to higher amounts of SFA and especially of C16:0, which affects the triacylglycerol structure and consequently has an effect on the thermal properties of milk fat (Tzompa-Sosa *et al.*, 2014). Tzompa-Sosa *et al.* (2014) reported that a modification in the triacylglycerol structure suggests that the distribution of FA could be triggered by differences in blood-derived fatty acids or by changes in the activity of enzymes related to fat synthesis that respond to the availability of FA for triacylglycerol synthesis.

Overall, the selection for higher fat yield and fat content would increase all underlying fatty acids in milk throughout lactation ( $r_g$  from 0.17 to 0.98). Likewise, there were positive correlations between milk protein content and FA, ranging from lower at the beginning ( $r_g$  from 0.06 to 0.57) to higher magnitude at the end of lactation ( $r_g$  from 0.56 to 0.68). On the other hand, the genetic correlations between protein yield and FA were negative ( $r_g$  from  $-0.11$  to  $-0.19$ ). Similar curves across lactation for milk FA with protein yield and content were also shown by Fleming *et al.* (2018) using a fifth-order Legendre polynomial. Nevertheless, they found weaker estimates which were close to zero for the association between protein content and LCFA and UFA in early lactation, as well as lower values in late lactation.

There was a strong genetic correlation estimated between milk FA from the middle to the end of lactation, which suggests that selection for one-time point will likely result in genetic gains for all lactation stages. In addition, individual FAs were strongly associated with the group to which they belonged, as seen by the average genetic correlation between C16:0 and SFA (0.98) and between C18:1 *cis*-9 and LCFA (0.91), and UFA (0.93). C18:1 *cis*-9 is the single UFA with the highest concentration in milk accounting for approximately 8 g/l whole milk (Haug *et al.*, 2007), which could explain the high association found. Genetic correlations between C18:1 *cis*-9 and C16:0 and SFA had approximately the same average estimates (i.e., 0.44 and 0.45, respectively). Therefore, selection in favor of C18:1 *cis*-9 is likely to yield similar results for these FAs. It seems that the genetic mechanism that drives de novo (i.e., half of C16:0) FA synthesis also drives the FA originating from other syntheses. Previous results were expected as well as an estimated high genetic correlation between C16:0 and SFA (0.98) because palmitic acid is a 16-carbon SFA and is most commonly found in animals (Loften *et al.*, 2014). Furthermore, similar genetic gains would have been expected for LCFA and UFA regarding selection on C18:1 *cis*-9 ( $r_g$  0.91 and 0.93, respectively).

The lowest genetic correlations were estimated between C18:1 *cis*-9 and C16:0, and between C18:1 *cis*-9 and SFA at the beginning of lactation. These estimates can be explained by the different origins of FA and its dynamic pattern influenced by lactation stage, energy balance and dietary composition. Milk fat is the main component determining energy expenditure for milk production in dairy cows, and most ruminant adipose tissue is stored as triglycerides, which comprise mainly C16:0, C18:0, and C18:1 *cis*-9 (Chilliard *et al.*, 2000; Gross *et al.*, 2011). In early lactation,

**Table 2.** Mean validation reliabilities ( $r^2$ ), regression coefficients ( $\hat{b}_1$ ) and their respective standard deviation for parentage average (PA) and genomic breeding value estimated assuming different\* scaling ( $\omega$ ) and weighting ( $\alpha$ ) factors for milk FA traits for the Walloon Holstein bulls.

Scenario	C16:0			C18:1 cis-9			LCFA			SFA			UFA		
	$r^2$	$\hat{b}_1$	$r^2$	$r^2$	$\hat{b}_1$	$r^2$	$r^2$	$\hat{b}_1$	$r^2$	$r^2$	$\hat{b}_1$	$r^2$	$r^2$	$\hat{b}_1$	
PA	0.06 (0.02)	0.54 (0.14)	0.03 (0.02)	0.03 (0.02)	0.35 (0.23)	0.03 (0.02)	0.03 (0.02)	0.43 (0.25)	0.10 (0.02)	0.10 (0.02)	0.69 (0.08)	0.04 (0.03)	0.04 (0.03)	0.42 (0.25)	
$\omega$ 1 $\alpha$ 0.95	0.10 (0.07)	0.34 (0.21)	0.04 (0.03)	0.04 (0.03)	0.13 (0.09)	0.04 (0.03)	0.04 (0.03)	0.19 (0.11)	0.16 (0.09)	0.16 (0.09)	0.46 (0.21)	0.05 (0.03)	0.05 (0.03)	0.28 (0.11)	
$\omega$ 1 $\alpha$ 0.8	0.10 (0.08)	0.39 (0.23)	0.03 (0.02)	0.03 (0.02)	0.17 (0.10)	0.04 (0.03)	0.04 (0.03)	0.17 (0.09)	0.16 (0.09)	0.16 (0.09)	0.51 (0.21)	0.04 (0.03)	0.04 (0.03)	0.23 (0.12)	
$\omega$ 1 $\alpha$ 0.7	0.10 (0.07)	0.42 (0.23)	0.03 (0.02)	0.03 (0.02)	0.19 (0.11)	0.04 (0.03)	0.04 (0.03)	0.19 (0.10)	0.16 (0.09)	0.16 (0.09)	0.54 (0.21)	0.04 (0.03)	0.04 (0.03)	0.25 (0.13)	
$\omega$ 1 $\alpha$ 0.6	0.10 (0.07)	0.45 (0.24)	0.03 (0.02)	0.03 (0.02)	0.21 (0.12)	0.04 (0.03)	0.04 (0.03)	0.22 (0.11)	0.16 (0.08)	0.16 (0.08)	0.58 (0.21)	0.05 (0.03)	0.05 (0.03)	0.28 (0.14)	
$\omega$ 0.9 $\alpha$ 0.95	0.10 (0.08)	0.37 (0.22)	0.03 (0.02)	0.03 (0.02)	0.14 (0.10)	0.04 (0.03)	0.04 (0.03)	0.16 (0.09)	0.16 (0.10)	0.16 (0.10)	0.51 (0.23)	0.04 (0.03)	0.04 (0.03)	0.20 (0.11)	
$\omega$ 0.8 $\alpha$ 0.95	0.10 (0.08)	0.40 (0.24)	0.03 (0.02)	0.03 (0.02)	0.15 (0.10)	0.04 (0.03)	0.04 (0.03)	0.17 (0.09)	0.17 (0.10)	0.17 (0.10)	0.56 (0.24)	0.04 (0.03)	0.04 (0.03)	0.22 (0.12)	
$\omega$ 0.7 $\alpha$ 0.95	0.11 (0.09)	0.43 (0.25)	0.03 (0.02)	0.03 (0.02)	0.16 (0.11)	0.04 (0.03)	0.04 (0.03)	0.19 (0.10)	0.17 (0.10)	0.17 (0.10)	0.60 (0.26)	0.04 (0.03)	0.04 (0.03)	0.23 (0.12)	
$\omega$ 0.6 $\alpha$ 0.95	0.11 (0.09)	0.47 (0.26)	0.03 (0.02)	0.03 (0.02)	0.17 (0.11)	0.04 (0.03)	0.04 (0.03)	0.20 (0.10)	0.17 (0.10)	0.17 (0.10)	0.65 (0.27)	0.04 (0.03)	0.04 (0.03)	0.25 (0.13)	
$\omega$ 0.6 $\alpha$ 0.6	0.11 (0.07)	0.63 (0.28)	0.04 (0.03)	0.04 (0.03)	0.32 (0.16)	0.05 (0.03)	0.05 (0.03)	0.34 (0.15)	0.17 (0.08)	0.17 (0.08)	0.81 (0.25)	0.05 (0.03)	0.05 (0.03)	0.41 (0.18)	
$\omega$ 0.7 $\alpha$ 0.6	0.11 (0.07)	0.60 (0.28)	0.04 (0.03)	0.04 (0.03)	0.30 (0.14)	0.05 (0.03)	0.05 (0.03)	0.32 (0.14)	0.17 (0.08)	0.17 (0.08)	0.77 (0.24)	0.05 (0.03)	0.05 (0.03)	0.38 (0.17)	

\* $\omega$ , scaling factor used for genomic relationship matrix ( $1.0G^{-1}$ );  $\omega$ , scaling factor used for pedigree relationship matrix ( $\omega A_{22}^{-1}$ );  $\alpha$  and  $\beta$ , weighting factors used for polygenic effect ( $ssGBLUP - \beta A_{22}$ )

dairy cows mobilize their body reserves to deliver the energy required for milk synthesis and secretion in the mammary gland. Thus, the FA in milk originates from both the mammary gland uptake of preformed FA from circulating blood (approximately 60%) and de novo synthesis within the mammary gland (approximately 40%) (Chilliard *et al.*, 2000). C18:1 cis-9 is the predominant FA in adipocytes and is primarily released through lipolysis during negative energy balance (NEB). When lipolysis is high, the FA composition of milk has a much higher proportion of C18:0 and C18:1 cis-9 (Barber *et al.*, 1997).

In early lactation, C16:0 originates primarily from mobilized body fat, and thereafter the cow achieves a positive energy balance and most of the C16:0 should now be produced within the mammary cells from acetate (Loften *et al.*, 2014). Furthermore, the lower genetic correlation found between C16:0 and LCFA in early lactation can be associated with this mobilization of body fat. In particular, LCFAs are derived from plasma and incorporated into milk, which inhibits de novo synthesis (part of C16:0) by the mammary gland. Selecting C18:1 cis-9 would result in a great increase in UFAs and LCFAs (genetic correlations above 0.90) in milk. Increased mobilization of lipids is associated with higher NEFA levels, which are particularly rich in LCFAs, especially C18:1 cis-9 (Jorjong *et al.*, 2014). Efforts to carry out practical improvements of these milk FAs are usually driven by their several roles firstly as possible biomarkers for early lactation metabolic disorders (Jorjong *et al.*, 2014), then from specifically nutritional effects to benefit consumers (Haug *et al.*, 2007) and finally from technological properties influencing phenomena such as oxidation and possible sensory changes (Hanuš *et al.*, 2018).

**Prediction reliability and bias**

Genomic predictions using optimal scaling and weighting factors in the ssGBLUP approach led to greater validation reliability and less bias compared to the traditional BLUP for most milk production traits. These findings are in agreement with previous studies in dairy cattle (Koivula *et al.*, 2015; Oliveira *et al.*, 2019). The ssGBLUP models using the optimal scaling and weighting factors improved the reliability by 0.03 for milk yield, 0.02 for fat and protein yield, and 0.07 and 0.08, respectively, for fat and protein content, which represent an increase ranging from 8.57% to 80%. The use of optimal factor scaling to combine  $G^{-1}$  and  $A_{22}^{-1}$  is required for a better model fit (increase  $r^2$  and decrease  $b_1$ ) and may better account for differences in the genetic architecture of each trait analyzed (Oliveira *et al.*, 2019). According to Gao *et al.* (2012), the bias of prediction tended to decrease with increasing polygenic weights in the  $G$  matrix and suggests that genetic markers number is not able to explain the total genetic variance. Thus, the polygenic effect would account for the residual genetic variance, which is not accounted for by using only genetic markers (Guarini *et al.*, 2018). As seen in our results, it becomes important to optimize the weighting factor  $\alpha$  used in the combination of the raw genomic and the pedigree relationship matrix.

Overall, differences in validation reliabilities  $r^2$  among different scenarios to build the  $H$  matrix were smaller than effects on regression coefficients  $b_1$ , which is consistent with Guarini *et al.* (2018). According to Koivula *et al.* (2015), the degree of inflation in GEBV is affected by the method used in construction of the  $H$  matrix. Decreasing  $\omega$  led to an increase in the regression coefficient and consequently decreased inflation, and according to Martini *et al.* (2018), this can be due to the fact that decreasing  $\omega$  tends to reduce the variance in the GEBV. Thus, the scaling

factors could be chosen to achieve smaller bias (degree of inflation), which was expected for young animals (Masuda *et al.*, 2016). Furthermore, decreasing  $\omega$  increases the importance of pedigree information in genomic prediction, and it is also dependent on the completeness of the pedigree. As reported by Misztal *et al.* (2017), ssGBLUP evaluations are inflated when the pedigree is deep but incomplete. The best  $\omega$  parameter assumed in this study was 0.70 (lower than 1.00) for milk production traits, which increases the importance of pedigree information on GEBV prediction. Likewise, Tsuruta *et al.* (2011) reported that smaller values for  $\omega$  (0.70) could be used to reduce the inflation of US Holstein genomic evaluations for young bulls without affecting accuracy.

For all milk FAs, the inclusion of genomic information based on the ssGBLUP approach also improved the reliabilities for young bulls in all scenarios evaluated. The combination of  $\omega$  equal to 0.60 associated with a polygenic effect  $\alpha$  equal to 0.60 yielded the least biased GEBVs predicted for milk FAs. These ssGBLUP models improved the accuracy by 0.05 for C16:0, 0.01 for C18:1 *cis*-9, 0.02 for LCFA, 0.07 for SFA, and 0.01 for UFA, representing an increase from 25% to 70%. Similarly, gains in prediction by using ssGBLUP for milk FA profiles predicted by MIR were also reported by Freitas *et al.* (2020) in dairy cattle. The prediction reliabilities were very low (ranging from 0.005 to 0.19) for C16:0 and C18:1 *cis*-9 predicted by gas chromatography and were presented by Gebreyesus *et al.* (2019) using GBLUP model in Chinese, Danish and Dutch Holstein cows. The effect of a small training population size and the lower heritability estimates may be the cause of the lower prediction reliability of GEBV. According to Guarini *et al.* (2018), predictions for lowly heritable traits benefit greatly from genomic information, especially when using the ssGBLUP approach. Improvement in genomic prediction may be achieved by increasing the numbers of both genotyped and phenotyped animals as well as using optimal scaling and weighting factors to maximize the observed accuracy of the GEBVs for milk FA in our population.

In conclusion, changes in milk production and FA traits can be achieved using genomic selection over days in milk. Selection for higher milk yield would decrease fat and protein content, as well as all FAs (C16:0, C18:1 *cis*-9, LCFA, SFA, and UFA). Improving the milk FA profile (especially based on C16:0 and SFA) seems to be an effective way to indirectly select for fat yield and fat content. The ssGBLUP approach yielded higher reliabilities than the traditional BLUP for young bulls. A less biased GEBV was found by choosing optimal scaling factors in the construction of the **H** matrix. Therefore, ssGBLUP based on RRM is feasible for the genomic prediction of milk production and FA traits in Walloon Holstein cattle.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0022029922000474>

**Acknowledgment.** The authors acknowledge the support of the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement; SPW-DGARNE) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech provided by the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the National Fund for Scientific Research (Brussels, Belgium) under Grants T.0095.19 (PDR ‘DEEPSELECT’) and J.0174.18 (CDR ‘PREDICT-2’) and the Consortium des Equipements de Calcul Intensif (CECI) of the Federation Wallonia-Brussels (Brussels, Belgium), funded by the National Fund for Scientific Research (Brussels, Belgium) funded under Grant 2.5020.11. The authors also gratefully

acknowledge the financial and technical support provided by the Walloon Breeders Association (AWE, Ciney, Belgium). The authors also acknowledge the support of the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), the CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and the Wallonia-Brussels-International (Brussels, Belgium).

## References

- Aguilar I, Misztal I, Johnson D L, Legarra A, Tsuruta S and Lawlor TJ (2010) Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science* **93**, 743–752.
- Aguilar I, Misztal I, Tsuruta S and Legarra A (2014) PREGSF90 – POSTGSF90 : Computational Tools for the Implementation of Single-step Genomic Selection and Genome-wide ... Proceedings, 10th World Congress of Genetics Applied to Livestock Production PREGSF90, 0–3.
- Barber MC, Clegg RA, Travers MT and Vernon RG (1997) Lipid metabolism in the lactating mammary gland. *Biochimica et Biophysica Acta – Lipids and Lipid Metabolism* **1347**, 101–126.
- Bastin C, Gengler N and Soyeurt H (2011) Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. *Journal of Dairy Science* **94**, 4152–4163.
- Chilliard Y, Ferlay A, Mansbridge RM and Doreau M (2000) Ruminant milk fat plasticity: nutritional control of saturated, polyunsaturated, trans and conjugated fatty acids. *Annales de Zootechnie* **49**, 181–205.
- Christensen OF and Lund MS (2010) Genomic relationship matrix when some animals are not genotyped Genomic prediction models. *Genetics Selection Evolution* **42**, 1–8.
- Colinet FG, Vandenplas J, Vanderick S, Hammami H, Mota RR, Gillon A, Hubin X, Bertozzi C and Gengler N (2017) Bayesian single-step genomic evaluations combining local and foreign information in Walloon Holsteins. *Animal* **12**, 898–905.
- Druet T, Jaffrézic F, Boichard D and Ducrocq V (2003) Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *Journal of Dairy Science* **86**, 2480–2490.
- Fleming A, Schenkel FS, Malchiodi F, Ali RA, Mallard B, Sargolzaei M, Jamrozik J, Johnston J and Miglior F (2018) Genetic correlations of mid-infrared-predicted milk fatty acid groups with milk production traits. *Journal of Dairy Science* **101**, 4295–4306.
- Freitas PHF, Oliveira HR, Silva FF, Fleming A, Miglior F, Schenkel FS and Brito LF (2020) Genomic analyses for predicted milk fatty acid composition throughout lactation in North American Holstein cattle. *Journal of Dairy Science* **103**, 6318–6331.
- Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS and Su G (2012) Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics Selection Evolution* **44**, 8.
- Gebreyesus G, Bovenhuis H, Lund MS, Poulsen NA, Sun D and Buitenhuis B (2019) Reliability of genomic prediction for milk fatty acid composition by using a multi-population reference and incorporating GWAS results. *Genetics Selection Evolution* **51**, 16.
- Gross J, van Dorland HA, Bruckmaier RM and Schwarz FJ (2011) Milk fatty acid profile related to energy balance in dairy cows. *Journal of Dairy Research* **78**, 479–488.
- Guarini AR, Lourenco DAL, Brito LF, Sargolzaei M, Baes CF, Miglior F, Misztal I and Schenkel FS (2018) Comparison of genomic predictions for lowly heritable traits using multi-step and single-step genomic best linear unbiased predictor in Holstein cattle. *Journal of Dairy Science* **101**, 8076–8086.
- Hanuš O, Samková E, Křížová L, Hasoňová L and Kala R (2018) Role of fatty acids in milk fat and the influence of selected factors on their variability-A review. *Molecules* **23**, 1636.
- Haug A, Høstmark AT and Harstad OM (2007) Bovine milk in human nutrition – a review. *Lipids in Health and Disease* **6**, 25.
- Jorjong S, van Kneegsel ATM, Verwaeren J, Lahoz MV, Bruckmaier RM, De Baets B, Kemp B and Fievez V (2014) Milk fatty acids as possible



- biomarkers to early diagnose elevated concentrations of blood plasma non-esterified fatty acids in dairy cows. *Journal of Dairy Science* **97**, 7054–7064.
- Koivula M, Strandén I, Pösö J, Aamand GP and Mäntysaari EA** (2015) Single-step genomic evaluation using multitrait random regression model and test-day data. *Journal of Dairy Science* **98**, 2775–2784.
- Legarra A, Christensen OF, Aguilar I and Misztal I** (2014) Single step, a general approach for genomic selection. *Livestock Science* **166**, 54–65.
- Loften JR, Linn JG, Drackley JK, Jenkins TC, Soderholm CG and Kertz AF** (2014) Invited review: palmitic and stearic acid metabolism in lactating dairy cows. *Journal of Dairy Science* **97**, 4661–4674.
- Lourenco D, Legarra A, Tsuruta S, Masuda Y, Aguilar I and Misztal I** (2020) Single-step genomic evaluations from theory to practice: using SNP chips and sequence data in BLUPF90. *Genes* **11**, 790.
- Mäntysaari E, Liu Z and VanRaden PM** (2010) Interbull validation test for genomic evaluations. *Interbull Bulletin* **51**, 17–22.
- Martini JWR, Schrauf MF, Garcia-Baccino CA, Pimentel ECG, Munilla S, Rogberg-Munóz A, Cantet RJC, Reimer C, Gao N, Wimmer V and Simianer H** (2018) The effect of the  $H - 1$  scaling factors  $\tau$  and  $\omega$  on the structure of  $H$  in the single-step procedure. *Genetics Selection Evolution* **50**, 16.
- Masuda Y, Misztal I, Tsuruta S, Legarra A, Aguilar I, Lourenco DAL, Fragomeni BO and Lawlor TJ** (2016) Implementation of genomic recursions in single-step genomic best linear unbiased predictor for US Holsteins with a large number of genotyped animals. *Journal of Dairy Science* **99**, 1968–1974.
- Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P and Baes CF** (2017) A 100-year review: identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science* **100**, 10251–10271.
- Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T and Lee DH** (2002) Blupf90 and Related Programs (Bgf90), 2001–2002.
- Misztal I, Bradford HL, Lourenco DAL, Tsuruta S, Masuda Y and Legarra A** (2017) Studies on inflation of GEBV in single-step GBLUP for type. *Interbull Bulletin*.
- Oliveira HR, Lourenco DAL, Masuda Y, Misztal I, Tsuruta S, Jamrozik J, Brito LF, Silva FF and Schenkel FS** (2019) Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. *Journal of Dairy Science* **102**, 2365–2377.
- Paiva JT, Mota RR, Lopes PS, Hammami H, Vanderick S, Oliveira HR, Veroneze R, Silva FF and Gengler N** (2022) Random regression test-day models to describe milk production and fatty acid traits in first lactation Walloon Holstein cows. *Journal of Animal Breeding and Genetics*. *Journal of Animal Breeding and Genetics* **139**, 398–413.
- Penasa M, Tiezzi F, Gottardo P, Cassandro M and De Marchi M** (2015) Genetics of milk fatty acid groups predicted during routine data recording in Holstein dairy cattle. *Livestock Science* **173**, 9–13.
- Tsuruta S, Misztal I, Aguilar I and Lawlor TJ** (2011) Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. *Journal of Dairy Science* **94**, 4198–4204.
- Tzompa-Sosa DA, van Aken GA, van Hooijdonk ACM and van Valenberg HJF** (2014) Influence of C16:0 and long-chain saturated fatty acids on normal variation of bovine milk fat triacylglycerol structure. *Journal of Dairy Science* **97**, 4542–4551.
- VanRaden PM** (2008) Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423.