ORIGINAL ARTICLE

Journal of
Animal Breeding and Genetics   **WILEY**

Genetic evaluation of gestation length in Italian Holstein breed

Ferdinando Galluzzo[1,2] | Giulio Visentin[2] | Johannes B. C. H. M. van Kaam[1] | Raffaella Finocchiaro[1] | Stefano Biffani[3] | Angela Costa[2] | Maurizio Marusi[1] | Martino Cassandro[1,4]

[1]Associazione Nazionale Allevatori della Razza Frisona, Bruna e Jersey Italiana (ANAFIBJ), Cremona, Italy

[2]Department of Veterinary Medical Sciences, Alma Mater Studiorum - University of Bologna, Ozzano dell'Emilia, Bologna, Italy

[3]Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, Milan, Italy

[4]Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Legnaro, Padova, Italy

Correspondence
Giulio Visentin, Department of Veterinary Medical Sciences, Alma Mater Studiorum - University of Bologna, Via Tolara di Sopra 50, Ozzano dell'Emilia, Bologna, 40064, Italy.
Email: giulio.visentin@unibo.it

Funding information
European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) - MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 - D.D. 1032 17/06/2022, CN00000022).

Abstract

Gestation length (GL) can potentially affect health and performance of both the dam and the newborn calf, and it is controlled by two genetic components, direct and maternal. This means that both the calf (direct effect) and the cow (maternal effect) genotypes contribute to determine GL and its variability. The aims of the present study were to estimate direct and maternal variance components of GL, develop a routine genetic evaluation of GL in Italian Holstein and evaluate potential (un)favourable associations with traits for which selection is undertaken in this population. A multiple-trait repeatability linear animal model was employed for the estimation of variance components considering GL in first and later parities as different traits. The posterior mean (PM) of heritability of the direct effect was 0.43 for first parity and 0.35 for later parities. The PM of heritability of the maternal effect was lower, being 0.08 for primiparae and 0.06 for pluriparae. The posterior standard deviation (PSD) of the heritability estimates was small, ranging from 0.001 to 0.005. The relationship of direct and maternal effects with important traits such as milk yield and fertility indicated that selecting for extreme GL, longer or shorter, may have negative consequences on several traits, suggesting that GL has an intermediate optimum in dairy cattle. In conclusion, this study reveals that selecting an intermediate GL in the Italian Holstein population is advisable. Although scarcely variable compared to other conventional traits for which Italian Holstein is selected, GL is heritable and a deeper knowledge can be useful for decision-making at the farm level.

KEYWORDS
dairy cattle, genomic selection, Gibbs sampling, pregnancy, SNPBLUP

1 | INTRODUCTION

Gestation length (GL) is defined as the interval from conception to calving and it is expressed in days. This trait is known to be genetically affected by both the calf's (direct effect, GL_d) and the cow's (maternal effect, GL_m) genotype (Eaglen et al., 2013; Norman et al., 2009, 2011) with an effect on the performance of the dam and the newborn calf (Eaglen et al., 2013). Indeed, the same study reported an unfavourable (i.e., negative) genetic correlation between GL_d and milk yield (MY) and protein yield (PY). The same authors reported also a positive genetic correlation between GL_m and number of services per pregnancy. As indicated in Norman et al. (2011), bulls' genetic proofs of GL are informative of the daughters' performance also for productive and functional traits, such as calving ease, stillbirth (SB) incidence, days open (DO), MY, PY, fat yield (FY), somatic cell score and survival. Considering the predicted transmitting ability, a non-linear relationship with optimum intermediate for GL was observed with all these traits with the exception of somatic cell score. In particular, bulls with intermediate merit for GL seemed to be the most favourable for the vast majority of the traits of interest. Jamrozik et al. (2005) reported an unfavourable genetic correlation between first parity GL_d and both calving ease (0.17) and SB (−0.11), suggesting that a longer GL is expected to be associated with a greater incidence of neonatal mortality and dystocia (DYS). At the phenotypic level, however, Meyer et al. (2001) reported a decrease in SB as GL increases. Other studies reported a non-linear relationship between GL and SB, with an intermediate GL as the optimum (Johanson & Berger, 2003; McGuirk et al., 1999; Niskanen & Juga, 1998; Philipsson, 1976a).

Genetic variation is a necessary condition to genetically improve a trait of interest in selected populations. The heritability of GL_d was estimated in different cattle populations (Bourdon & Brinks, 1982; Haile-Mariam & Pryce, 2019; Johanson et al., 2011; McGuirk et al., 1999; Philipsson, 1976a) and falls within the range of 0.28–0.60. Although heritable, the maternal counterpart is generally lower (Haile-Mariam & Pryce, 2019; Johanson et al., 2011), that is GL_m is reported to be between 0.04 and 0.08. These results may confirm that calving is an event where cow and calf genotypes both play a key role and that parturition is triggered by a pre-partum strong increase in foetal corticosteroid concentrations (Hunter et al., 1977; Thorburn et al., 1977). This may explain why the direct effect of the calf, GL_d, has normally the greatest heritability.

As nowadays the importance of reducing direct and indirect costs has become crucial in dairy farming, a deep investigation into the possibility of selecting an optimal GL is advisable. In line with this, it becomes essential to estimate the effect of GL on other economically important traits in Holstein cattle.

Hence, the objectives of the present study were: (i) to estimate variance components of GL, (ii) to develop a genetic and genomic evaluation for this trait to be implemented in the Italian Holstein population and (iii) to quantify the association between genetic components (GL_d and GL_m) of GL and other traits undergoing routine evaluation.

2 | MATERIALS AND METHODS

2.1 | Data and data editing

Data employed in the current research originated from the Italian National Breeders Association (AIA, Rome, Italy) and consisted of 24,705,045 calving and services events of Holstein cows. For GL calculation, the successful insemination was considered the last one available for a given cow. Such insemination was restricted to be within 240–305 days before the subsequent calving date. From the original dataset, only calvings from 1997 onwards and with a GL between 260 and 305 days were kept and only records from animals with both parents known were used. Furthermore, records from not purebred inseminations and twin calvings were excluded. Regarding parity, only observations from cows up to 5th calving were analysed and the cow's age at parity was restricted between 18 and 40 months for primiparae, between 33 and 55 months for secondiparae, and between 43 and 97 months for the remaining ones. In order to have an effective comparison between contemporary groups, herd-year-season of conception (HYS) levels with less than 10 observations were excluded. The pedigree provided by the National Association of Holstein, Brown and Jersey Breed was traced back up to four generations.

2.2 | Statistical model

The genetic evaluation was based on a multiple trait repeatability linear animal model, considering GL phenotypes in first and later parities as different but correlated traits. To select the fixed effects to be included, significance was preliminarily tested using the *statsmodels* package available for Python (Seabold & Perktold, 2010).

The defined model for GL in first parity cows is described as follows (Model 1):

$$y_{ijklmnopq} = \mu + \text{HYS}_i + (\text{MOB} - \text{YOB})_{jk} + (\text{SEX} - \text{YOB})_{lm} + \text{AGEC}_n + a_o + d_p + e_{ijklmnopq}$$

with $y_{ijklmnopq}$ as the phenotypic observation of GL expressed in days; u as overall intercept; HYS_i as the fixed effect of the ith HYS; $(\text{MOB} - \text{YOB})_{jk}$ as the fixed effect of the jth month of birth (MOB) within the kth year of birth (YOB); $(\text{SEX} - \text{YOB})_{lm}$ as the fixed effect of the lth sex of the newborn calf within the mth year of birth; AGEC_n as the fixed effect of age at first calving n of the dam (age at calving between 18 and 20 months, inclusive, were grouped into a unique class); a_o as the random additive genetic effect of the oth calf; d_p as the random additive genetic effect of the pth dam; and $e_{ijklmnopq}$ as the random residual of the qth observation. In Model 1, the fixed effect of sex was considered within year of birth to account for the progressive increase in the use of sexed semen in Holstein heifers in the last years. For calvings originating from embryo transfer, the direct effect was linked to the donor, while the maternal to the recipient.

The defined model for GL in older cows is described as follows (Model 2):

$$y_{ijklmnop} = \mu + \text{HYS}_i + (\text{MOB} - \text{YOB})_{jk} + \text{SEX}_l + P_m + mpe_n + a_o + d_n + e_{ijklmnop}$$

Compared to Model 1, Model 2 did not account for the interaction between year of birth and sex, and different parities were present; therefore, SEX_l was the fixed effect of the lth sex of the newborn calf and P_m was the fixed effect of the mth class of dam parity order (1, 2, 3, 4+: parity 4 and 5 were grouped into a common class). Random effects were mpe_n as the permanent environment effect of the nth dam because of the presence of across-lactations repeated measures, a_o as the additive genetic of the oth calf, d_n as the additive genetic of the nth dam, and $e_{ijklmnop}$ as the residual of the pth observation.

2.3 | Genetic and genomic analyses

2.3.1 | Estimation of genetic parameters

Variance components estimation was performed with the Gibbs sampler THRGIBBS1F90 (Misztal et al., 2002), by using the aforementioned described models for first- and later-parities. The analysis was performed on the entire dataset, with 120,000 iterations, a burn-in of 45,000 and a thinning rate of 10. Post-Gibbs analysis was performed using the software POSTGIBBSF90 (Misztal et al., 2002) using the remaining 7500 samples. Convergence was assessed by visual inspection and the POSTGIBBSF90 diagnostics available included the Posterior Standard Deviation (PSD) as uncertainty measure.

For Model 1, the direct (h_d^2) and maternal (h_m^2) heritabilities were calculated as described below:

$$h_d^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2 + \sigma_e^2}, \text{ and } h_m^2 = \frac{\sigma_m^2}{\sigma_a^2 + \sigma_m^2 + \sigma_e^2}$$

with σ_a^2 as the direct additive genetic variance, σ_m^2 as the maternal additive genetic variance and σ_e^2 as the residual variance.

For Model 2, h_d^2 and h_m^2 were calculated as follows:

$$h_d^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2 + \sigma_{mpe}^2 + \sigma_e^2}, \text{ and } h_m^2 = \frac{\sigma_m^2}{\sigma_a^2 + \sigma_m^2 + \sigma_{mpe}^2 + \sigma_e^2},$$

with σ_{mpe}^2 as the maternal permanent environment effect variance.

Finally, the additive genetic coefficient of variation (CV_a, %) was calculated for all of the four genetic effects (first and later parities GL_d and first and later parities GL_m) by dividing the additive genetic standard deviation by the phenotypic mean for each trait and multiplying it by 100.

2.4 | Prediction and validation of EBV

The estimated variance components were used to predict estimated breeding values (EBV) with BLUP using the software MiX99 (MiX99 Development Team, 2019). The BLUP solutions were standardized to the mean and standard deviation of the genetic base, which included cows born between 2014 and 2016. Then, EBV were expressed on a scale with mean of 100 and standard deviation of 5 so that proofs >100 indicate individuals transmitting longer GL. Generated EBV from model 1 were subsequently validated to determine the bias and accuracy of predictions of $GL_{d \text{ and }} GL_m$. LR method proposed by Legarra and Reverter (2018) was applied. Briefly, EBV were obtained from both a full (with all data until June 2022) and a reduced (partial) dataset. In the reduced dataset phenotypic observations of the last 4 years were excluded. The BLUP solutions of both models were then standardized to the mean and standard deviations of the solutions of cows born between 2008 and 2010, which were considered as the genetic base. For GL_d the focal individuals were 246 bulls born between 2018 and 2020 with 10 to 15 daughters in the full dataset and no daughters in the partial one. For GL_m, the focal group was composed of 200 bulls born

between 2016 and 2020; the filtering criteria for daughters were the same as above. Four statistics were computed on the focal groups:

(i) Bias ($\mu_{w,p}$) as $\frac{\hat{\mu}_p - \hat{\mu}_w}{\sigma_p}$, with $\hat{\mu}_p$ as the average of the solutions' vector in the partial dataset, $\hat{\mu}_w$ as the average of the solutions' vector in the full dataset and σ_p as the solutions' standard deviation in the partial dataset;

(ii) Dispersion ($b_{w,p}$) as $\frac{Cov(\hat{u}_p, \hat{u}_w)}{Var(b_p)}$, with $\hat{u}_p$ and $\hat{u}_w$ as the solutions' vectors in the partial and full dataset, respectively, Cov denotes a covariance, and Var denotes a variance;

(iii) Pearson correlation coefficient ($\rho_{w,p}$) between $\hat{u}_p$ and $\hat{u}_w$; and

(iv) Reliability ($\widehat{rel}$) as $\frac{Cov(\hat{u}_p, \hat{u}_w)}{(1-\overline{F})\hat{\sigma}_u^2}$, with $\overline{F}$ being the average inbreeding coefficient of individuals in the focal group and $\hat{\sigma}_u^2$ as the additive genetic variance.

In order to obtain a measure of uncertainty for the four statistics, bootstrap method was applied on 1000 samples with replacement of the focal group with the *bootstrap* module of the *SciPy* library of the Python programming language (Virtanen et al., 2020).

2.4.1 | Prediction and validation of direct genomic values

The prediction and validation of direct genomic values (DGV) for GL_d and GL_m was performed as described in Finocchiaro et al. (2012) and Galluzzo et al. (2022). Briefly, estimated de-regressed proofs (EDP) generated from previously calculated EBV (Degano et al., 2016) were obtained from the full and the reduced dataset described above. Bulls' genotypes ($n=8642$) generated from different DNA chips were imputed to a panel containing 69,952 single nucleotide polymorphisms (SNP) through the PedImpute software (Nicolazzi et al., 2013) with an allele error rate of 0.0012. A conventional quality control was applied to SNP data, considering the following criteria: call-rate ≥0.95, minor allele frequency ≥0.02 and Hardy–Weinberg equilibrium *p*-value ≥0.005. The genomic evaluation, based on the SNPblup method, was performed on the reduced dataset with EDP of training bulls used as pseudo-phenotypes. Generated SNP solutions (i.e., allelic substitution effects) were used to compute the DGV of validation bulls ($n=374$), with daughters in the full dataset but not in the reduced one.

The SNPblup model was as follows:

$$y = 1\mu + Mg + e$$

with y as the EDP vector, μ as overall intercept, M as the $n \times m$ matrix of genomic information, with n as the number of individuals and m as the number of SNP, g as the random SNP effects vector and e as the vector of residuals.

Finally, EDP in the full dataset were linearly regressed against bulls DGV estimated in the reduced run. The coefficient of determination, calculated as the squared Pearson correlation coefficient from the previous regression, was then used as the metric to validate genomic prediction. When both EBV and DGV were available, genomically enhanced breeding values (GEBV) were calculated as follows:

$$GEBV = \frac{EDC_c * EBV + EDC_g * DGV}{EDC_c + EDC_g},$$

where EDC_c is the conventional effective daughters contribution (Fikse & Banos, 2001), and EDC_g is the genomic effective daughters contribution (Patry & Ducrocq, 2011). The whole process of validation, for both EBV and DGV, is graphically described in Figure 1.

## 2.5 | Relationships with other traits

Approximate genetic correlations were calculated as the Pearson correlation coefficient between all available GEBV, including within and across parities GL_d and GL_m, on a dataset composed of 73,659 genotyped cows born after 2015. In order to evaluate the relationship between cows' performances and mates' GL_d or sires' GL_m EBVs, mixed models were fitted using Echidna Mixed Model software (Gilmour, 2020). Four phenotypes were analysed: first lactation milk solids production (MS1, regarded as fat + protein yield), lifetime milk solids production (MSY), first lactation calving difficulty score (CD1, 1–3 scale with higher values meaning higher difficulty; farmer's subjective score) and the first lactation SB (SB1). Bulls were stratified into seven classes based on the rounded standard deviation units of their standardized EBV (−3: 83–87, −2: 88–92, −1: 93–97, 0: 98–102, 1: 103–107, 2: 108–112 and 3: 113–117). For the direct effect, the mate's EBV group was fitted as fixed effect, while for the maternal effect the father's EBV group was used.

The model used to obtain the least squares means to investigate the effect of bull GL_d on mates' performances was:

$$Y_{ijkl} = HY_i + AGEC_j + dGLg_k + e_{ijkl}$$

with Y_{ijkl} as the cow's phenotypic observation of the trait analysed, HY_i as the fixed effect of the ith herd-year of first calving, $AGEC_j$ as the fixed effect of the jth cow's age at first

—WILEY⊥

FIGURE 1 Steps for the validation of (a) EBV and (b) DGV[1] using both a full dataset and a reduced[2] dataset. [1]Bulls selected for validation (n = 374) did not have daughters in the reduced dataset. [2]Gestation length phenotypes recorded from 2018 onwards were masked. [Colour figure can be viewed at wileyonlinelibrary. com]
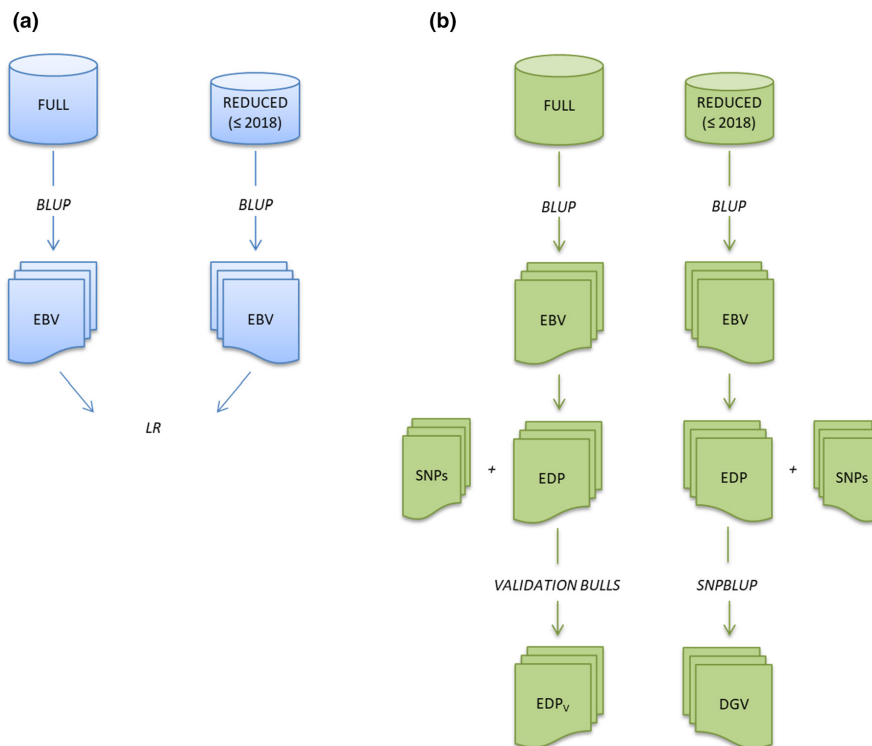
calving, $dGLg_k$ as the fixed effect of kth class of sire GL_d EBV and e_{ijkl} as the residual of the lth observation.

The model used to investigate the effect of sire GL_m on daughters' performances through the least squares means was:

$$Y_{ijkl} = HY_i + AGEC_j + mGLg_k + e_{ijkl}$$

with Y_{ijkl} as the cow's phenotypic observation of the trait analysed, HY_i as the fixed effect of the ith herd-year of first calving, $AGEC_j$ as the fixed effect of the jth cow's age at first calving, $mGLg_k$ as the fixed effect of the kth class of sire GL_m EBV and e_{ijkl} as the residual of the lth observation.

3 | RESULTS

3.1 | Descriptive statistics

The median of GL before editing was 279 days, with the first and the third quartile at 275 and 285 days. After the removal of outliers and edits, the final dataset consisted of 8,796,487 GL records of single calves, delivered by 4,741,508 dams (1.86 observations per dam on average) on 8274 herds (1063 observations per herd on average) divided in 149,793 contemporary groups (58.72 observations per group, on average). 40% of all observations belonged to first parity dams, while 60% to later parities. The frequency of male calves was 49%. Descriptive statistics of GL by calf sex and parity order are depicted in Table 1.

TABLE 1 Descriptive statistics of the gestation length (GL) expressed in days after edits.

Trait	n	Mean	Standard deviation
GL, days	8,796,487	278.75	6.03
GL male calves, days	4,256,824	279.28	6.13
GL female calves, days	4,539,663	278.26	5.88
GL first-parity cows, days	3,489,606	278.02	6.01
GL later-parity cows, days	5,306,881	279.25	5.99

The average GL was 278.75 ± 6.03 days and male calves tended to be delivered later (+1.02 days) than females, with an average GL of 279.28 ± 6.13 versus 278.26 ± 5.88 days. Heifers tended to deliver earlier (278.02 ± 6.01 days) than pluriparous cows (279.25 ± 5.99 days), with an average difference of 1.23 days. For all groups in Table 1, minimum and maximum values for GL were 260 and 305 days, respectively, due to the filtering criteria described in the Material and Methods section.

3.2 | Genetic and genomic analyses

3.2.1 | Estimation of genetic parameters

The results of the variance components' estimation are listed in Table 2. Post-Gibbs analysis gave a PM of 0.43

Trait	CV_a, %	GL_d Parity 1	GL_d Parity >1	GL_m Parity 1	GL_m Parity >1
GL_d, parity 1	1.38	**0.43 (0.005)**	0.92 (0.006)	−0.50 (0.01)	−0.29 (0.01)
GL_d, parity >1	1.24		**0.35 (0.004)**	−0.29 (0.008)	−0.22 (0.009)
GL_m, parity 1	0.61			**0.08 (0.002)**	0.85 (0.01)
GL_m, parity >1	0.50				**0.06 (0.001)**

TABLE 2 Coefficient of additive genetic variation (CV_a), posterior mean of the heritability (posterior standard deviation) in the diagonal (in bold) and genetic correlation (above diagonal) for direct (GL_d) and maternal (GL_m) gestation length in first and later parities.
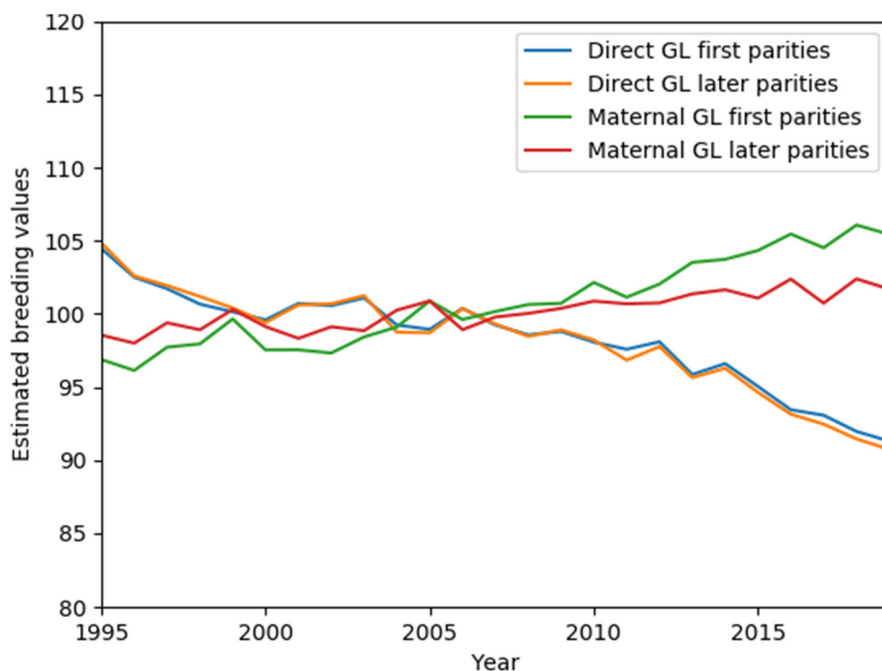

FIGURE 2 Trend of bulls estimated breeding values for direct and maternal gestation length (GL) in first and later parity across year of birth. [Colour figure can be viewed at wileyonlinelibrary.com]

with a PSD of 0.005 and the highest probability density of 5%–95% of 0.420–0.442 for first-parity GL_d heritability and 0.08 (PSD = 0.002; HPD$_{5-95\%}$ = 0.080, 0.086) for first-parity GL_m one. For later parities, the PM for GL_d and GL_m heritability was 0.35 (PSD = 0.004; HPD$_{5-95\%}$ = 0.339, 0.355) and 0.06 (PSD = 0.001; HPD$_{5-95\%}$ = 0.056, 0.060), respectively. The genetic correlation between the GL_d and GL_m was −0.50 (PSD = 0.01; HPD$_{5-95\%}$ = −0.516, −0.491) for first parities and −0.22 (PSD = 0.009; HPD$_{5-95\%}$ = −0.236, −0.204) for later parities. The within-trait genetic correlation between the two parity levels was 0.92 (PSD = 0.006; HPD$_{5-95\%}$ = 0.909, 0.928) for GL_d and 0.85 (PSD = 0.01; HPD$_{5-95\%}$ = 0.838, 0.866) for GL_m. A negative genetic correlation of −0.29 (PSD = 0.01; HPD$_{5-95\%}$ = −0.310, −0.275) was estimated between first-parity GL_d and later-parities GL_m and between later-parities GL_d and first-parity GL_m (PSD = 0.008; HPD$_{5-95\%}$ = −0.313, −0.273).

Finally, irrespectively of the parity groups considered, the CV_a was greater in GL_d compared to GL_m (Table 2): the low CV_a for both direct (1.38% and 1.24% for first and later parities respectively) and maternal effects (0.61% for first parity and 0.50% for later parities), weakens the possibility to select for GL.

3.2.2 | Genetic trend

The genetic trend by bulls' year of birth for GL_d and GL_m is represented in Figure 2 and reflects the negative genetic correlation between these components. While direct GL has a decreasing trend, the maternal one is slightly increasing, especially for first parities over the last 20 years. The genetic trend of GL_d of first mimics that of later parities, due to their strong but not unity genetic correlation. This strong decreasing trend in GL_d is evident at the phenotypic level too: in the last 20 years, the mean GL of the Italian Holstein population has decreased by nearly 2 days (unpublished data).

3.2.3 | Validation of EBV and DGV

The results of the LR method for first-parity GL_d and GL_m on the focal group are reported in Table 3. Regarding GL_d, the bias was −0.177 genetic standard deviations, with a bootstrap confidence interval 95% (BCI) ranging from −0.182 and −0.176. Dispersion was 0.90 (BCI = 0.895, 0.905) while the correlation and reliability were 0.586

TABLE 3 LR validation results of pedigree indexes for first parity direct (GL_d) and maternal (GL_m) gestation length; the 95% bootstrap confidence interval is reported in parentheses.

Trait	Bias	Dispersion	Correlation	Reliability
GL_d	−0.177 (−0.182, −0.176)	0.900 (0.895, 0.905)	0.586 (0.584, 0.589)	0.259 (0.257, 0.261)
GL_m	−0.101 (−0.106, −0.956)	0.952 (0.948, 0.957)	0.670 (0.668, 0.673)	0.154 (0.153, 0.155)

FIGURE 3 Approximate genetic correlations[1] between the genomically enhanced breeding values (GEBV) of direct (GL_d; blue bars) and maternal (GL_m; orange bars) and the GEBV available from routine genetic and genomic evaluation of Italian Holstein cattle. [1]Pearson correlation coefficient between sets of GEBV of 73,659 genotyped cows born after 2015. [Colour figure can be viewed at wileyonlinelibrary.com]
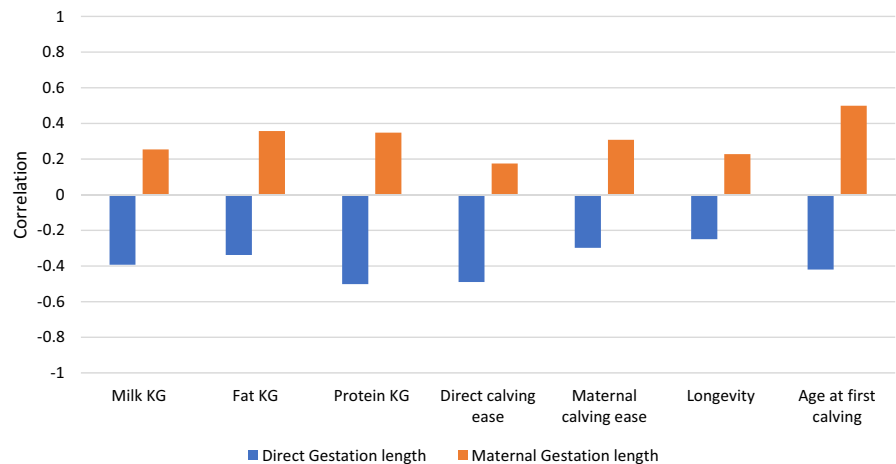

FIGURE 4 Least squares means of mate's direct gestation length (GL) EBV group with cow's phenotypic performances: (a) milk solids yield in first lactation, (b) lifetime milk solids yield, (c) first lactation calving difficulty score, and (d) first lactation stillbirth. Asterisks indicate statistical difference at $p < 0.05$ compared to the least squares mean of class 0. [Colour figure can be viewed at wileyonlinelibrary.com]
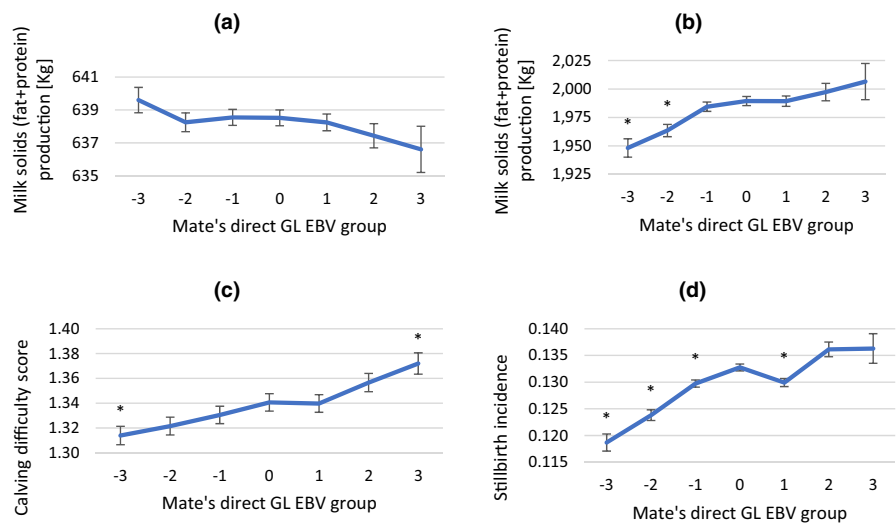
(BCI = 0.584, 0.589) and 0.259 (BCI = 0.257, 0.261), respectively. The bias of GL_m was −0.101 (BCI = −0.106, −0.956), its dispersion was 0.952 (BCI = 0.948, 0.957) and the correlation and reliability averaged 0.67 and 0.154, respectively, with BCI for the first ranging from 0.668 to 0.673 and from 0.153 to 0.155 for the latter.

The result of the DGV validation for reliability was 0.516 for GL_d and 0.223 for GL_m (data not shown).

3.3 | Relationships with other traits

Approximate genetic correlations between GL_d and GL_m to other traits included in the routine genetic evaluation are presented in Figure 3. For GL_d, the correlation

with MY (−0.40), PY (−0.50) and fat yield (FY, −0.34) were moderately negative. Regarding functional traits, a negative correlation was estimated with direct longevity (−0.25) and direct (−0.40) and maternal DYS (−0.30). The correlation with age at first calving (AFC) was negative (−0.42). For GL_m, a positive relationship was quantified with MY (0.25), FY (0.36) and PY (0.35). As far as it concerns functional traits, results from the present study indicated a slightly positive correlation with longevity (0.23) and a positive correlation with DYS direct (0.18) and DYS maternal (0.31). Finally, a moderate positive correlation with AFC (0.50).

An overview of performance for MS1, MSY, CD1, and SB1 estimated for each class of bulls' EBV is presented in Figure 4 (GL_d) and Figure 5 (GL_m). A decreasing trend for
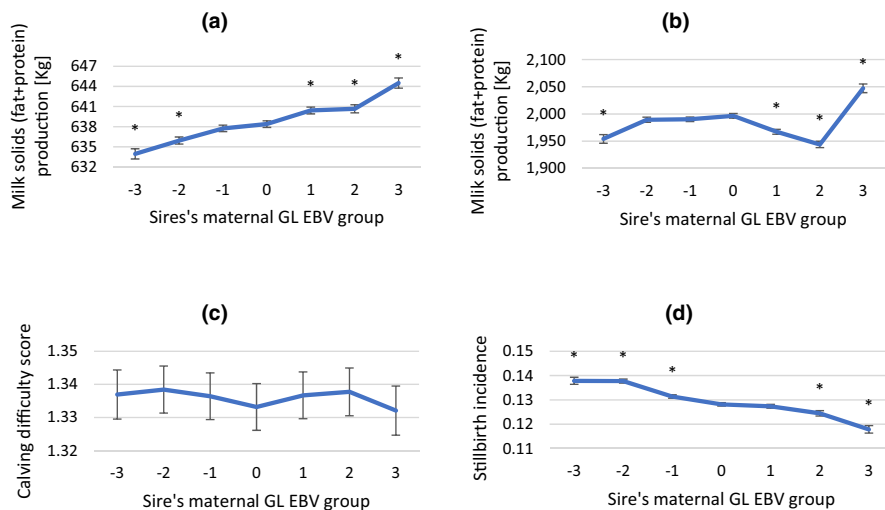
FIGURE 5 Least squares means of sire's maternal gestation length (GL) EBV group with daughter's phenotypic performances: (a) milk solids yield in first lactation, (b) lifetime milk solids yield, (c) first lactation calving difficulty score, and (d) first lactation stillbirth. Asterisks indicate statistical difference at $p < 0.05$ compared to the least squares mean of class 0. [Colour figure can be viewed at wileyonlinelibrary.com]

MS1 across classes of bulls GL_d was observed, although differences among the estimated least squares means were not significant. However, considering MSY, such a trend across bull GL_d classes turned positive and a significant difference was quantified between class −3 (i.e., proof between 83 and 87) and 0 (i.e., proof between 98 and 102) and between −2 (i.e., proof between 88 to 92) and 0. Both CD1 and SB1 increased concurrently with greater bulls GL_d proofs. The effect of sire GL_m EBV on daughters' performance for the same aforementioned traits is presented in Figure 5. Least squares means for MS1 increased with GL_m classes, and an opposite trend exists for SB1, with significant differences between opposite classes in both instances. An erratic trend was estimated for MSY and CD1 with desirable values for bull's GL_m proof between 98 and 102.

4 | DISCUSSION

4.1 | Scope for genetic selection

Heritability estimates of GL_d from the present study corroborates with literature estimates. For example, McGuirk et al. (1999), in UK Holstein Friesian using a sire model, reported an estimate of 0.45 (standard error, 0.02), while GL_d heritability of 0.51 (0.05) was quantified in Johanson et al. (2011), applying an animal model in US Holstein adjusting for the parity of the dam. In Australian Holstein, heritabilities in first (0.28) and later parity (0.36) GL_d were estimated using a sire-maternal grandsire model (Haile-Mariam & Pryce, 2019). In contrast, in UK Holstein Friesian, Eaglen et al. (2012) reported heritability values of 0.57 and 0.41 for first parity and later parity GL_d, respectively. Regarding GL_m, literature heritability estimates range from 0.04 (Haile-Mariam & Pryce, 2019) to 0.08 (Johanson et al., 2011), in

agreement with the results of the present study. Non-zero heritability of a trait is a necessary criterion in order to be considered for genetic selection (Miglior et al., 2017). Estimates of the present study confirm the possibility of selecting GL, especially for GL_d. Moreover, the higher heritability of the direct component compared to the maternal one suggests that the effect of the calf on GL is stronger than that of the mother, confirming that the role of the corticosteroids secretion of the fetus is crucial for the triggering of parturition (Hunter et al., 1977; Thorburn et al., 1977). The small coefficient of genetic variation reported in the current study may suggest that direct selection on these animal features may have limited results: however, a small change in phenotypic GL is known to affect DYS and SB (Philipsson, 1976b), probably due to the high foetal growth rate occurring in the last month of gestation (Meijering, 1984). Moreover, phenotypic GL was observed to affect lactation performance (Norman et al., 2011; Vieira-Neto et al., 2017), likely due to the high mammary gland parenchymal tissue growth rate at the end of gestation (Davis, 2017).

The genetic correlation between first parity and later parities for both GL_d (0.92) and GL_m (0.85) means that, for both components, first and later parities should be treated as different traits in a genetic evaluation. On the other hand, the genetic correlation between GL_d and GL_m for both first parities (−0.50) and later parities (−0.22) suggests that a female calf which genetically tends to shorten the gestation while in utero, tends to extend it when pregnant. These negative correlations are in accordance with findings reported for both beef and dairy cattle breeds (Cervantes et al., 2010; Haile-Mariam & Pryce, 2019; Jeyaruban et al., 2016; Johanson et al., 2011; Philipsson, 1976a). In particular, Haile-Mariam and Pryce (2019) reported a genetic correlation of −0.31 between GL_d and GL_m for both first and later parity, which is slightly lower but comparable with the

results of the present study. These results may be explained by a natural selection mechanism that guarantees a balance between the calf's and dam's development (Philipsson, 1976a); furthermore, female calves secreting high levels of corticosteroid to trigger the dam's partum may be more tolerant to foetal corticosteroids once they are pregnant. However, a deeper investigation to validate this hypothesis is needed. The validation of bulls' pedigree indexes (PIs) gave a greater reliability for GL_d compared to GL_m, which can be explained by the differences in the heritability estimates of the two components. The results for bias for both direct and maternal GL suggest that the estimates are nearly unbiased while a slight overestimation was highlighted by both dispersion coefficients. Finally, a slightly higher correlation between full and reduced run was detected for GL_m compared to GL_d.

As for the PIs, the reliability of DGVs was higher for the direct (0.516) than maternal component (0.223) and, in general, noticeably higher than the reliability of PIs.

4.2 | Relationships to other traits

Regarding GL_d, the approximate genetic correlations with other traits highlighted inverse relationship with MY, PY and FY, meaning that bulls transmitting longer GL tend to transmit a smaller genetic potential for milk-related performances. Eaglen et al. (2013) reported GL_d to be genetically correlated with both MY (−0.19) and PY (−0.22) in UK Holstein-Friesian primiparous cows. Their results, although slightly different from those of the present study, confirm the inverse relationship between these traits. Regarding functional traits, the present study reveals that longer GL is associated with lower longevity and higher risk of DYS. For longevity, a non-significant positive correlation (0.09) with GL_d was reported in UK Holstein-Friesian (Eaglen et al., 2013). In the same population, an unfavourable (positive) genetic correlation (0.34) has been estimated between GL and DYS (McGuirk et al., 1999). A study on Canadian Holstein revealed an unfavourable positive genetic correlation of 0.17 between GL_d and DYS directly evaluated in primiparous cows (Jamrozik et al., 2005). Johanson et al. (2011) identified an unfavourable positive correlation of 0.38 between GL_d and DYS direct and a slightly favourable negative one with DYS maternal (−0.06). Regarding fertility, the results of the present study suggests that bulls with a high GEBV for GL_d tend to produce a progeny that calves later than the average.

For the maternal component, the GL_m, a positive relationship was detected with MY, FY and PY. A previous study, in contrast, indicated a non-significant relationship between maternal GL and production traits (Eaglen

et al., 2013). As concerns the functional traits, no correlation was found with longevity but a positive relationship was observed with DYS, both direct and maternal. Eaglen et al. (2013) reported a negative, although not significant, genetic correlation of −0.23 between GL_m and longevity. In the case of DYS, literature results are contrasting, with no consensus about the relationship between GL and CE, also partly due to the population studied and statistical models adopted. A positive and unfavourable genetic correlation (0.19) between CE and GL_m was observed in Canadian Holstein (Jamrozik et al., 2005), while Johanson et al. (2011) reported in US Holstein a favourable negative genetic correlation of −0.24 with DYS direct, which is in accordance with the results of the present study. Johanson et al. (2011) reported an unfavourable positive genetic correlation of 0.25 of maternal DYS with maternal GL.

Regarding the relationship of the mate's direct GL EBV with first lactation milk content production of the cow, a slightly negative trend in production as the EBV increase was observed, although not significant. In contrast, analysing lifetime milk contents production, the trend is slightly positive, with a lifetime production of cows mated with bulls with lower GL EBV significantly lower than the others. Focusing on calving traits, although the majority of coefficients were insignificant, a positive trend was observed with calving difficulty score, meaning that calving difficulty increases as gestation length does; analysing SB incidence, the present study reveals a positive trend too, with shorter GL resulting in a lower risk of SB.

Regarding sire's maternal GL EBV, a significant favourable positive relationship was observed with first lactation milk contents production. In the case of the lifetime performance, instead, a semi-intermediate optimum (slightly shifted towards shorter GL) was detected, although a significantly higher production level was observed for the daughters of bulls present in the +3 standard deviations of GL EBV group, maybe due to the smaller sample size. Analysing calving traits, an intermediate optimum was identified for the relationship of sire's GL EBV group with calving difficulty score, although the differences between coefficients were not significant due to the high standard errors. Focusing on the SB incidence, a significant favourable (negative) relationship was observed, meaning that daughters of bulls with greater GL_m EBVs tend to deliver more vital newborns, that is, with a reduced neonatal mortality in the first 48 hours of life. These results seem to confirm that bulls' genetic potential for GL_d and GL_m plays a role on calving and subsequent lactation performances of cows, even if the effect of management and genetic potential for the traits that directly regulate the phenotypes

of interest are stronger. Moreover, this study strengthens the hypothesis that, in the absence of specific management needs, neither a shortening nor a prolongation of GL should be pursued at population level.

5 | CONCLUSIONS

The present study wanted to deepen the knowledge of the genetic aspects of GL in Italian Holstein. Despite the phenotypic and genetic variability of GL being rather small due to biological constraints, the results of the present study reveal that GL has a heritability different from zero, confirming that potentially there could be room to select for this trait. This study also indicates that maintaining GL in an intermediate optimum is advisable and does not cause undesired responses in other traits of interest, such as milk yield, solids yield, or fertility. Furthermore, given that in Italy there is no seasonality and calvings are equally distributed along the year, there is no need to shorten GL for management purposes. Of the two genetic variance components evaluated, direct and maternal, the direct one is predominant: this difference strengthens the hypothesis that the calf genotype's effect is crucial at calving and triggers parturition. The genetic correlation between the direct and the maternal components highlighted, for both traits, the antagonistic relationship between the effect of an animal while in utero and while pregnant. The genetic correlations between parities indicated that first and later parities should be treated as different traits in the genetic evaluation for GL. Both the validation of PIs and DGVs confirmed that the identified model is stable enough and suitable for routine genetic and genomic evaluation in this population. The results of this study partially clarify the link between the genetic, phenotypic and physiological aspects of GL and lead to the implementation of a routine genetic evaluation of GL in the Italian Holstein breed, providing a new tool for decision-making at the farm level.

ORCID
Ferdinando Galluzzo https://orcid.org/0000-0003-2301-2656
Giulio Visentin https://orcid.org/0000-0003-0869-5516
Johannes B. C. H. M. van Kaam https://orcid.org/0000-0002-2592-2461
Raffaella Finocchiaro https://orcid.org/0000-0002-9058-9992
Stefano Biffani https://orcid.org/0000-0001-5559-3630
Angela Costa https://orcid.org/0000-0001-5353-8988
Maurizio Marusi https://orcid.org/0000-0002-2359-3633
Martino Cassandro https://orcid.org/0000-0002-8709-2870

REFERENCES

Bourdon, R. M., & Brinks, J. S. (1982). Genetic, environmental and phenotypic relationships among gestation length, birth weight, growth traits and age at first calving in beef cattle. *Journal of Animal Science*, *55*, 543–553. https://doi.org/10.2527/jas1982.553543x

Cervantes, I., Gutiérrez, J. P., Fernández, I., & Goyache, F. (2010). Genetic relationships among calving ease, gestation length, and calf survival to weaning in the Asturiana de los Valles beef cattle breed. *Journal of Animal Science*, *88*, 96–101. https://doi.org/10.2527/jas.2009-2066

Davis, S. R. (2017). Triennial lactation symposium/bolfa: Mammary growth during pregnancy and lactation and its relationship with milk yield. *Journal of Animal Science*, *95*, 5675–5688. https://doi.org/10.2527/jas2017.1733

Degano, L., Jansen, G., Finocchiaro, R., Rossoni, A., & Vicario, D. (2016). *Hybrid one-step genomic evaluation system for the Italian Simmental breed*. Interbull Bulletin NO. 50. Puerto Varas, Chile, October 24–28.

Eaglen, S. A. E., Coffey, M. P., Woolliams, J. A., & Wall, E. (2012). Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. *Genetics, Selection, Evolution*, *44*, 23. https://doi.org/10.1186/1297-9686-44-23

Eaglen, S. A. E., Coffey, M. P., Woolliams, J. A., & Wall, E. (2013). Direct and maternal genetic relationships between calving ease, gestation length, milk production, fertility, type, and lifespan of Holstein-Friesian primiparous cows. *Journal of Dairy Science*, *96*, 4015–4025. https://doi.org/10.3168/jds.2012-6229

Fikse, W. F., & Banos, G. (2001). Weighting factors of sire daughter information in international genetic evaluations. *Journal of Dairy Science*, *84*, 1759–1767. https://doi.org/10.3168/jds.S0022-0302(01)74611-5

Finocchiaro, R., van Kaam, J. B. C. H. M., & Biffani, S. (2012). *The genomic selection system in Italian Holstein* (p. 130). EAAP.

Galluzzo, F., van Kaam, J.-T., Finocchiaro, R., Marusi, M., Tsuruta, S., & Cassandro, M. (2022). Estimation of milkability breeding values and variance components for Italian Holstein. *JDS Communications*, 3, 180–184. https://doi.org/10.3168/jdsc.2021-0167

Gilmour, A. R. (2020). Echidna mixed model software. https://www.echidnamms.org/

Haile-Mariam, M., & Pryce, J. E. (2019). Genetic evaluation of gestation length and its use in managing calving patterns. *Journal of Dairy Science*, 102, 476–487. https://doi.org/10.3168/jds.2018-14981

Hunter, J. T., Fairclough, R. J., Peterson, A. J., & Welch, R. A. S. (1977). Foetal and maternal hormonal changes preceding normal bovine parturition. *Acta Endocrinologica*, 84, 653–662. https://doi.org/10.1530/acta.0.0840653

Jamrozik, J., Fatehi, J., Kistemaker, G. J., & Schaeffer, L. R. (2005). Estimates of genetic parameters for Canadian Holstein female reproduction traits. *Journal of Dairy Science*, 88, 2199–2208. https://doi.org/10.3168/jds.S0022-0302(05)72895-2

Jeyaruban, M. G., Johnston, D. J., Tier, B., & Graser, H. U. (2016). Genetic parameters for calving difficulty using complex genetic models in five beef breeds in Australia. *Animal Production Science*, 56, 927–933. https://doi.org/10.1071/AN14571

Johanson, J. M., & Berger, P. J. (2003). Birth weight as a predictor of calving ease and perinatal mortality in Holstein cattle. *Journal of Dairy Science*, 86, 3745–3755. https://doi.org/10.3168/jds.S0022-0302(03)73981-2

Johanson, J. M., Berger, P. J., Tsuruta, S., & Misztal, I. (2011). A Bayesian threshold-linear model evaluation of perinatal mortality, dystocia, birth weight, and gestation length in a Holstein herd1. *Journal of Dairy Science*, 94, 450–460. https://doi.org/10.3168/jds.2009-2992

Legarra, A., & Reverter, A. (2018). Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genetics, Selection, Evolution*, 50, 53. https://doi.org/10.1186/s12711-018-0426-6

MiX99 Development Team. (2019). *MiX99: A software package for solving large mixed model equations*. Release XI/2019. Natural Resources Institute Finland (Luke), Jokioinen, Finland. http://www.luke.fi/mix99

McGuirk, B. J., Going, I., & Gilmour, A. R. (1999). The genetic evaluation of UK Holstein Friesian sires for calving ease and related traits. *Animal Science*, 68, 413–422. https://doi.org/10.1017/S1357729800050414

Meijering, A. (1984). Dystocia and stillbirth in cattle – A review of causes, relations and implications. *Livestock Production Science*, 11, 143–177. https://doi.org/10.1016/0301-6226(84)90057-5

Meyer, C. L., Berger, P. J., Koehler, K. J., Thompson, J. R., & Sattler, C. G. (2001). Phenotypic trends in incidence of stillbirth for Holsteins in the United States. *Journal of Dairy Science*, 84, 515–523. https://doi.org/10.3168/jds.S0022-0302(01)74502-X

Miglior, F., Fleming, A., Malchiodi, F., Brito, L. F., Martin, P., & Baes, C. F. (2017). A 100-year review: Identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science*, 100, 10251–10271. https://doi.org/10.3168/jds.2017-12968

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). *BLUPF90 and related programs (BGF90). Commun.*

No. 28-07. Page in 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France.

Nicolazzi, E. L., Biffani, S., & Jansen, G. (2013). Short communication: Imputing genotypes using PedImpute fast algorithm combining pedigree and population information. *Journal of Dairy Science*, 96, 2649–2653. https://doi.org/10.3168/jds.2012-6062

Niskanen, S., & Juga, J. (1998). Calving difficulties and calf mortality in Finnish dairy cattle population. *Interbull Bulletin*, 18, 17–20.

Norman, H. D., Wright, J. R., Kuhn, M. T., Hubbard, S. M., Cole, J. B., & VanRaden, P. M. (2009). Genetic and environmental factors that affect gestation length in dairy cattle. *Journal of Dairy Science*, 92, 2259–2269. https://doi.org/10.3168/jds.2007-0982

Norman, H. D., Wright, J. R., & Miller, R. H. (2011). Potential consequences of selection to change gestation length on performance of Holstein cows. *Journal of Dairy Science*, 94, 1005–1010. https://doi.org/10.3168/jds.2010-3732

Patry, C., & Ducrocq, V. (2011). Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. *Genetics, Selection, Evolution*, 43, 30. https://doi.org/10.1186/1297-9686-43-30

Philipsson, J. (1976a). Studies on calving difficulty, stillbirth and associated factors in Swedish cattle breeds. *Acta Agriculturae Scandinavica*, 26, 230–234. https://doi.org/10.1080/00015127609435092

Philipsson, J. (1976b). Studies on calving difficulty, stillbirth and associated factors in Swedish cattle breeds: III. Genetic parameters. *Acta Agriculturae Scandinavica*, 26, 211–220. https://doi.org/10.1080/00015127609435090

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. In *Proceedings of the 9th Python in Science Conference* (pp. 92–96). https://doi.org/10.25080/Majora-92bf1922-011

Thorburn, G. D., Challis, J. R. C., & Currie, W. B. (1977). Control of parturition in domestic animals. *Biology of Reproduction*, 16, 18–27. https://doi.org/10.1095/biolreprod16.1.18

Vieira-Neto, A., Galvão, K. N., Thatcher, W. W., & Santos, J. E. P. (2017). Association among gestation length and health, production, and reproduction in Holstein cows and implications for their offspring. *Journal of Dairy Science*, 100, 3166–3181. https://doi.org/10.3168/jds.2016-11867

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2