J. Dairy Sci. TBC
https://doi.org/10.3168/jds.2025-26876

Comparing genetic architecture of mid-infrared-predicted energy balance, a novel energy deficiency score, and several biomarkers

Hongqing Hu,[1] Sébastien Franceschini,[1] Pauline Lemal,[1] Hadi Atashi,[1,2] Clément Grelet,[3] Yansen Chen,[1] Katrien Wijnrocx,[1] Soyeurt Hélène,[1] and Nicolas Gengler[1]*
[1]TERRA Teaching and Research Center, Gembloux Agro-Bio Tech (ULiège-GxABT), University of Liège, 5030 Gembloux, Belgium
[2]Department of Animal Science, Shiraz University, 71441-13131 Shiraz, Iran
[3]Walloon Agricultural Research Center (CRA-W), 5030 Gembloux, Belgium

ABSTRACT

Negative energy balance (NEB) during early lactation is a critical physiological challenge in high-producing dairy cows, affecting both their health and production performance. The objectives of this study were: (1) to compare the genetic architecture of logit-transformed predicted NEB (LPNEB), a logit-transformed novel energy deficiency score (LEDS), 15 biomarkers, and 3 production traits using SNP-based genomic correlation analysis; (2) to extend this study to a chromosomal level to identify specific genomic regions involved in the regulation of energy metabolism; and (3) to compare the independent contributions of 8 traits to the underlying genetic architecture of LPNEB and LEDS. The SNP effects estimated from single-trait models can be used to quickly calculate genomic correlations for 20 traits. The results indicate strong genomic correlations between LPNEB and LEDS, as well as with key metabolic biomarkers, particularly blood nonesterified fatty acids (NEFA), highlighting their importance in energy metabolism. Furthermore, NEFA was a strong independent contributor to both LPNEB and LEDS. Chromosome regions located on BTA19 and BTA25 were identified as potentially associated with NEB. By combining genomic correlation and contribution analyses, this study provides valuable insights into the genetic basis of NEB and related traits in dairy cows.

Key words: genomic correlation, genetic architecture, independent contribution

INTRODUCTION

Negative energy balance (**NEB**) is a common metabolic state in high-yielding dairy cows, especially during early lactation, when energy intake does not meet the physiological demands of maintenance and milk production (Churakov et al., 2021). Although the short-term mobilization of body reserves to sustain lactation represents a normal adaptive mechanism, prolonged NEB may have detrimental effects on health (Hammon et al., 2006), reproduction performance (Wathes et al., 2007), and immune competence (Esposito et al., 2014; Mekuriaw, 2023). These physiological challenges contribute to increased veterinary costs, labor, and culling, ultimately reducing animal welfare and herd profitability (Heringstad et al., 2000; Gohary et al., 2016; Miglior et al., 2017).

Understanding the genetic architecture underlying NEB-related traits is critical for developing effective genomic selection strategies aimed at improving metabolic resilience. Previous studies have reported moderate h^2 estimates for NEB and related metabolic traits, generally ranging from 0.15 to 0.40 (Becker et al., 2021; Hu et al., 2025), suggesting meaningful potential for genetic improvement.

Traditionally, NEB status is indirectly evaluated by directly measuring biomarkers through invasive blood sampling, which is not suitable for large-scale implementation also due to animal welfare concerns (Macrae et al., 2019). Furthermore, most studies rely on only 1 or 2 biomarkers, such as nonesterified fatty acids (**NEFA**) and BHB, which may not fully capture the genetic complexity of NEB (Knob et al., 2021). These limitations highlight the value of an alternative phenotyping procedure that is effective, noninvasive, and scalable. Recent advances in mid-infrared (**MIR**) spectrometry have provided a promising, large-scale, and noninvasive approach for phenotyping traits related to energy balance in dairy cows (Ho et al., 2019; Smith et al., 2019). By analyzing routine milk samples, MIR enables the prediction of physiological indicators, such as NEFA, BHB, and IGF-1 (Grelet et al., 2019; Aernouts et al., 2020; Benedet et al., 2020).

To explore the potential application of combining multiple MIR-based indicators, Franceschini et al. (2022) used clustering analysis and proposed a novel composite trait, the energy deficiency score (**EDS**),

Received May 3, 2025.
Accepted August 29, 2025.
*Corresponding author: nicolas.gengler@uliege.be

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-25. Nonstandard abbreviations are available in the Notes.

which was derived from 27 MIR-predicted metabolic features, such as milk BHB and oleic acid (C18:1 *cis*-9). The EDS-predicted model achieved a very high accuracy of 0.99 (Franceschini et al., 2024), which is much higher than that of the blood NEFA prediction model (coefficient of determination is 0.39; Grelet et al., 2019). For genetic analysis purposes, EDS was logit-transformed and referred to as logit-transformed novel energy deficiency score (**LEDS**). In the study by Hu et al. (2025), LEDS showed strong genetic correlations with logit-transformed predicted NEB (**LPNEB**, 0.85), blood BHB (0.69), and milk C18:1 *cis*-9 (0.81), further supporting its potential as a biologically relevant and genetically informative trait related to energy metabolism during early lactation. Moreover, the h^2 of LEDS was estimated at 0.18 (Hu et al., 2025), indicating a moderate genetic basis and supporting its potential utility in genetic selection programs.

Direct biomarkers such as blood NEFA, blood BHB, blood glucose, and milk fatty acids (**FA**) reflect physiological responses to NEB and are routinely used in research and diagnostics. Nevertheless, their genetic interrelationships with MIR-derived traits remain poorly understood. This lack is even more true for the traits LPNEB and LEDS (Hu et al., 2025). It is unclear whether these indicators share common genomic foundations or reflect distinct regulatory mechanisms, particularly when comparing the associations between LPNEB and the biomarkers of interest with the corresponding associations between LEDS and the same biomarkers.

Therefore, the objectives of this study done in first-parity Holstein cows were: (1) to compare the genetic architecture of LPNEB, LEDS, 15 biomarkers, and 3 production traits using SNP-based genomic correlation analysis; (2) to extend this study to a chromosomal level to identify specific genomic regions involved in the regulation of energy metabolism; and (3) to compare the independent contributions of 8 traits to the underlying genetic architecture of LPNEB and LEDS in first-parity Holstein cows. The findings from this study may enhance our understanding of the genetic basis of NEB and its related traits.

MATERIALS AND METHODS

Data Collection and Editing

Phenotypic Data. All phenotypic records were obtained from official milk data recording in the Walloon Region of Belgium between 2012 and 2019. The milk samples were analyzed by MIR spectrometry by commercial instruments FT+2, FT7, and FT6000 spectrometers (Foss, Hillerød, Denmark) and a Standard Lactoscope FT-MIR automatic (Delta Instruments, Drachten, the Netherlands) standardized into a common format (Grelet et al., 2015). The references and parameters of the MIR prediction equations used for 19 MIR-predicted traits are described by Hu et al. (2025) in their Table 1. The LPNEB was derived from the predicted energy balance (**PEB**) as follows:

$$\mathrm{LPNEB} = \log_{10}\frac{\mathrm{PNEB}}{1-\mathrm{PNEB}},$$

where $\mathrm{PNEB} = 1 - \frac{\mathrm{PEB} - \mathrm{PEB}_{\mathrm{minimum}}}{\mathrm{PEB}_{\mathrm{maximun}} - \mathrm{PEB}_{\mathrm{minimum}}}$, and the predictive model for PEB was developed using partial least squares regression. The coefficient of determination of PEB model in the cross-validation set was 0.43 (Grelet et al., 2017).

The LEDS was derived from the EDS as follows:

$$\mathrm{LEDS} = \log_{10}\frac{\mathrm{EDS}}{1-\mathrm{EDS}},$$

where predictive model for EDS was developed using partial least squares discriminant analysis. In the validation set, the model achieved an overall accuracy of 0.99, with a sensitivity of 0.95 and a specificity of 0.92 (Franceschini et al., 2024).

The data used consisted of test-day records of traits including LPNEB, LEDS, 15 biomarkers with $\log_{10}$-transformed blood BHB (**LB_BHB**), blood NEFA, $\log_{10}$-transformed blood IGF-1 (**LIGF-1**), blood glucose (**GLU**), $\log_{10}$-transformed milk BHB (**LM_BHB**), milk citrate (**CIT**), $\log_{10}$-transformed milk acetone (**LACE**), milk decanoic acid **(C10:0)**, milk myristic acid (**C14:0**), milk palmitic acid (**C16:0**), milk stearic acid (**C18:0**), milk short-chain FA (SCFA), milk medium-chain FA (**MCFA**), milk long-chain FA (**LCFA**), milk C18:1 *cis*-9, and 3 production traits: milk fat percentage (**FP**), milk protein percentage (**PP**), and milk yield (**MY**). Milk MIR spectra, FP, and PP were generated by MIR commercial instruments mentioned previously. The MY, FP, and PP were limited from 3 to 99 kg/d, 1% to 9%, and 1% to 7%, respectively (ICAR, 2022).

Global H distances between each standardized MIR spectrum in this study and the MIR spectra in the calibration dataset of the different traits were calculated (Whitfield et al., 1987). We only kept MIR records with Global H ≤3. Furthermore, all predicted traits were constrained within a DIM range of 5 to 50, as the predictive models for certain biomarkers are exclusively applicable within this DIM range (e.g., LPNEB, B_BHB).

The final dataset comprises 30,634 records on 25,287 primiparous Holstein cows distributed in 508 herds. The distribution of DIM for the retained records is presented

in Supplemental Figure S1, file 1, see Notes. The pedigree was traced back to 4 generations, resulting in 74,662 animals (5,017 males).

Genotypic Data. Genotypic data on either phenotyped animals or those included in the pedigree (3,757 animals, comprising 1,150 males and 2,607 females) were used. Individuals were genotyped using the BovineSNP50 Beadchip versions 1 to 3 and EuroG MD (SI) version 9 (Illumina, San Diego, CA). The SNPs common across the 4 chips were retained for analysis. Nonmapped SNP, SNPs located on sex chromosomes, and triallelic SNPs were excluded from the analysis. Only SNPs with a minimum GenCall Score of 0.15 and a minimum GenTrain Score of 0.55 were retained (Atashi et al., 2024). The genotypes were then imputed to the BovineHD using Flmpute V2.2 software (Sargolzaei et al., 2014), SNPs with Mendelian conflicts, as well as those with a minor allele frequency less than 5% were excluded from the analysis. The difference between observed and expected heterozygosity was estimated, and SNPs with a difference greater than 0.15 were removed (Wiggans et al., 2009). Ultimately, a total of 566,170 SNPs located on 29 BTA were retained for genomic analysis.

Variance Component Estimation. Variances of 20 traits were estimated using single-trait repeatability model. The following model was used:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Mm} + \mathbf{Cc} + \mathbf{Dd} + \mathbf{Qq} + \mathbf{Wp} + \mathbf{Za} + \mathbf{e},$$

where **y** is the vector of observations (LPNEB, LEDS, 15 biomarkers, and 3 production traits); **h** is the vector of fixed effect for herd × calving year; **m** is the vector of fixed effect of calving age (10 classes: 23–24, 25, 26, 27, 28, 29, 30, 31–32, 33 and 34, and 35–37 mo); **c** is the vector of fixed effect for calving month (12 class month); **d** is the fixed vector of DIM classes (for classes: DIM 5–6, 7–8, 9–10, and 11–50); **q** is the fixed regression coefficient vector of standardized DIM, and its quadratic; **p** is the vector of the random permanent environmental effects; **a** is the vector of random additive genetic effects; and **e** is the vector of residual effects. Additionally, **H**, **M**, **C**, **D**, **Q**, **W**, and **Z** are incidence matrices. Both classes and regressions were used to model the very rapid changes of certain biomarkers at the beginning of the lactation. When calculating the relationship between animals, we used the **H** matrix, which combined pedigree- (**A**) and genomic- (**G**) based relationships into one matrix. The inverse of **H** is as follows (Aguilar et al., 2010):

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

where **A** is the numerator relationship matrix based on the pedigree; $\mathbf{A}_{22}$ is the numerator relationship matrix based on the pedigree for genotyped animals; and **G** is the weighted genomic relationship matrix obtained using the following function:

$$\mathbf{G} = \mathbf{G}^* \times 0.95 + \mathbf{A}_{22} \times 0.05, \quad [2]$$

where $\mathbf{G}^*$ represents the genomic relationship matrix as defined by the first method of VanRaden (2008).

Computations were performed using the BLUPF90 family of programs (Misztal et al., 2018). The variance components and parameters for the 20 traits were estimated by AI-REML. The formulas used for calculating h^2 and repeatability (r) are as follows:

$$h^2 = \frac{\tilde{A}_a^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2} \quad [3]$$

$$r = \frac{\sigma_a^2 + \sigma_p^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2}, \quad [4]$$

where $\tilde{A}_a^2$, $\tilde{A}_p^2$, and $\tilde{A}_e^2$ represent the additive genetic, permanent environmental, and residual variances, respectively.

SNP Effect Estimation. The SNP effects ($\hat{\boldsymbol{\beta}}$) for each of the 20 traits were estimated using the POSTGSF90 software (version 1.73; Aguilar et al., 2014). The estimated method of the $\hat{\boldsymbol{\beta}}$ was the same as described by Wang et al. (2012) but without iteration. The formula to obtain $\hat{\boldsymbol{\beta}}$ for each trait is as follows:

$$\hat{\boldsymbol{\beta}} = \mathbf{D}\mathbf{Z}_\mathbf{g}^{'}\left(\mathbf{Z}_\mathbf{g}\mathbf{D}\mathbf{Z}_\mathbf{g}^{'}\right)^{-1}\hat{\mathbf{u}}, \quad [5]$$

where **D** = I, meaning all SNPs are given equal weight; $\mathbf{Z}_\mathbf{g}$ is an incidence matrix of genotyped for each SNP; and $\hat{\mathbf{u}}$ is a vector of GEBV for each trait.

Estimation of SNP-based Genomic Correlations. This analysis used the elements of the vector $\hat{\boldsymbol{\beta}}$ of SNP effect estimates obtained from the above step, along with the corresponding vector **p** of allele frequencies, to construct 20 × 20 trait correlation matrices across all chromosomes and at the level of each individual chromosome. These matrices included LPNEB, LEDS, and 18 other traits. The weighted covariance between trait i and trait j was calculated as:

$$Cov_{ij} = \sum_{k=1}^{n} 2\mathbf{p}_\mathbf{k}\left(1 - \mathbf{p}_\mathbf{k}\right)\boldsymbol{\beta}_{\mathbf{ik}}\boldsymbol{\beta}_{\mathbf{jk}}, \quad [6]$$

where $\mathbf{p_k}$ is the allele frequency of the *k*th SNP, $\boldsymbol{\beta}_{ik}$ and $\boldsymbol{\beta}_{jk}$ represent the SNP effect estimates for trait i and trait j, respectively, and *n* is the total number of SNPs on the chromosome. This approach follows the framework described by Gardner and Latta (2007), who explored how shared QTL contribute to genetic correlations between traits. While our analysis is based on genome-wide SNP effects derived from the single-trait ssGWAS, the logic of weighting by allele frequency to compute trait-level covariance is conceptually similar. The weighted genetic variance for trait i was calculated as:

$$Var_i = \sum_{k=1}^{n} 2\mathbf{p_k}\left(1-\mathbf{p_k}\right)\boldsymbol{\beta}_{ik}^2. \quad [7]$$

Using these values, the SNP-based genomic correlation between traits i and j was derived as:

$$r_{ij} = \frac{Cov_{ij}}{\sqrt{Var_i \cdot Var_j}} = \frac{\sum 2\mathbf{p}\left(1-\mathbf{p}\right)\cdot\boldsymbol{\beta}_{ij}}{\sqrt{\sum 2\mathbf{p}\left(1-\mathbf{p}\right)\cdot\boldsymbol{\beta}_i^2 \cdot \sum 2\mathbf{p}\left(1-\mathbf{p}\right)\cdot\boldsymbol{\beta}_j^2}}. \quad [8]$$

This analysis was conducted in R (version 4.4.2) using a custom function, and the output was a symmetric 20 × 20 correlation matrix representing the linear genomic relationships among traits.

The SE of the SNP-based genomic correlation was approximated using the following formula:

$$\mathrm{SE}\left(\mathrm{r}\right) \approx \sqrt{\frac{\left(1-r^2\right)^2}{\mathrm{n}-2}},$$

where r is the SNP-based genomic correlation and n = 556,170 is the number of SNP used in the analysis.

To test whether chromosome-wise correlations significantly differed from the genome-wide value, Fisher's r-to-z transformation was used. For each r, the z-score was calculated as

$$\mathrm{z} = \frac{1}{2}\ln\left(\frac{1+\mathrm{r}}{1-\mathrm{r}}\right).$$

The significance of the difference between the chromosome-specific correlation $\mathrm{r}_{\mathrm{BTA}}$ and the genome-wide correlation $\mathrm{r}_{\mathrm{All}}$ was assessed using the following test statistic:

$$\mathrm{z}_{\mathrm{diff}} = \frac{\mathrm{z}_{\mathrm{BTA}} - \mathrm{z}_{\mathrm{All}}}{\sqrt{\frac{1}{\mathrm{n}_{\mathrm{BTA}}-3} + \frac{1}{\mathrm{n}_{\mathrm{All}}-3}}},$$

where $\mathrm{n}_{\mathrm{BTA}}$ and $\mathrm{n}_{\mathrm{All}}$ denote the numbers of SNPs used to estimate the correlations for each chromosome and the whole genome, respectively.

The corresponding *P*-values were calculated for each chromosome. To adjust for multiple testing, the Benjamini-Hochberg false discovery rate (**FDR**) procedure was applied, and chromosome-wise correlations with FDR-adjusted *P*-values less than 0.05 were considered significant.

For each chromosome, the mean absolute genetic correlation (**MeanAbsCorr**) between LPNEB (or LEDS) and the other 19 traits was calculated using the following formula:

$$\mathrm{MeanAbsCorr}_{\mathrm{BTA}} = \frac{\sum_{i=1}^{19}\left|r_{\mathrm{BTA,i}}\right|}{19},$$

where $r_{\mathrm{BTA,i}}$ represents the genomic correlation between LPNEB (or LEDS) and the i-th trait on a given chromosome.

The chromosome with the largest MeanAbsCorr was identified, and its value was statistically compared with the genome-wide mean absolute correlation (ALL_BTA) using Fisher's r-to-z transformation and FDR correction (same as the above).

Estimation of the Independent Contributions of 8 Traits to LPNEB and LEDS. A total of 8 traits LB_BHB, NEFA, LIGF-1, LM_BHB, CIT, C10:0, PP, and MY were retained from the original set of 20 studied traits using the variance inflation factor (**VIF**) method (Miles, 2005) applied to the overall genomic correlations across all chromosomes. Traits with VIF values greater than 10 were iteratively removed to reduce multicollinearity. This filtering process ensured that the final set of traits used in the independent contribution analysis of LPNEB and LEDS had acceptable collinearity interpretability.

To quantify the independent contributions of 8 traits to LPNEB and LEDS, we used a selection index-based approach using SNP-based genomic correlations. This analysis was performed at both the whole-genome level (**All_BTA**) and across each of the 29 autosomes (BTA1–BTA29). For each analysis, the target trait (LPNEB or LEDS) was excluded from the correlation matrix, and the vector of partial regression coefficients was estimated using

$$\mathbf{w} = \mathbf{r}_{\mathrm{target}}\mathbf{C}^{-1}, \quad [9]$$

where $\mathbf{r}_{target}$ is the line vector of correlations between the target trait and other selected traits, and $\mathbf{C}^{-1}$ is the inverse of the submatrix of the correlation matrix among the other traits.

The total independent contribution of all traits was then calculated as

$$\text{total contribution} = \mathbf{wCw}^{t}, \quad [10]$$

where $\mathbf{w} = \mathbf{r}_{target}\mathbf{C}^{-1}$ and represents the vector of independent standardized contributions of the traits to the target trait in the presence of all other traits. This measure reflects the overall genetic signal explained by the selected traits after accounting for trait intercorrelations. This procedure was conducted separately for SNP effects on each chromosome to evaluate the specific contributions of the 8 traits at the chromosome level. Finally, to compare the independent standardized contributions of each trait within chromosomes, the contributions were expressed relative to each other by dividing each value by the sum of the absolute values of all trait contributions on the same chromosome i:

$$\text{relative contribution}_{i} = \frac{\mathbf{w}_{i}}{\sum_{j=1}^{n}\left|\mathbf{w}_{j}\right|}. \quad [11]$$

This approach allowed for the relative importance of each trait to be compared on the same scale within each chromosome. All matrix operations were conducted using base R software (version 4.4.2).

RESULTS AND DISCUSSION

Genetic Parameters

The h^2 and r estimates of the 20 traits are presented in Figure 1. Estimates obtained using the single-trait repeatability model were highly comparable to those derived from the multitrait repeatability model in our previous study (Hu et al., 2025). Most traits exhibited moderate h^2 (0.12–0.31), except for citrate, which had a higher estimate of 0.40. This suggests that citrate has a stronger genetic basis compared with the other analyzed traits.

Previous studies have reported h^2 estimates for energy balance ranging from 0.03 to 0.29, which are consistent with the estimates obtained in the present study (Buttchereit et al., 2011; Spurlock et al., 2012; Becker et al., 2021). The h^2 for the 4 predicted blood biomarkers (LB_BHB, NEFA, LIGF-1, and GLU) demonstrated strong concordance with previously reported values in the literature (Oikonomou et al., 2008; Benedet et al., 2020), supporting the reliability of these MIR-derived predictions. Similarly, the h^2 of LM_BHB, LACE, and CIT were consistent with those reported in previous studies (Koeck et al., 2014; Mehtiö et al., 2020; Chen et al., 2024). The h^2 of the 8 milk biomarkers and 3 production traits in this study are generally consistent with the range reported previously (Bastin et al., 2011; Benedet et al., 2020; Paiva et al., 2022). However, Atashi et al. (2023), using a random regression test-day model and data from 5 to 365 DIM, reported higher h^2 estimates for FA traits in first-parity Belgian Blue cows. This difference is likely due to differences in lactation stage and breed, as the present study focused on Holstein cows.

The r estimates for the 20 traits in this study ranged from 0.22 to 0.40, except for MY, which had an r of 0.58. Most traits exhibited moderate r, indicating that their variation is largely influenced by permanent environmental factors or genetic effects. In contrast, the higher r of MY suggests greater consistency across measurements, with less influence from environmental fluctuations.

Genomic Correlations Estimated Using SNP Effects for All Chromosomes

We estimated the genomic correlations among 20 traits using SNP effects derived from single-trait repeatability models (Figure 2). The values above the diagonal are based on the mentioned SNP-based estimates, while the values below the diagonal are from our previous 20-trait model (Hu et al., 2025). The results demonstrate a high degree of consistency between the 2 modeling approaches. The estimates from both models ranged from −0.8 and −0.79 (GLU and LB_BHB) to 0.97 (C16:0 and MCFA), with a minimal difference of only 0.01. Under the conditions of this study, the results of single-trait model provided reliable and consistent estimates of genetic correlations from multitrait model. Interestingly, the single-trait model based on SNP effects is computationally simpler and more suitable for large-scale multitrait genomic analyses.

This study found a strong genomic correlation (0.81) between LPNEB and LEDS, indicating a very similar genetic architecture. Both LPNEB and LEDS demonstrated strong genomic correlations, with absolute values greater than 0.5, with multiple biomarkers, including LB_BHB, NEFA, LIGF-1, LACE, C18:1 *cis*-9, and LCFA. The NEFA showed the highest genomic correlation with both LPNEB (0.80) and LEDS (0.81), confirming that it is a useful and early indicator of NEB, as it reflects how cows mobilize body fat and adjust their metabolism (Billa et al., 2020; Mehtiö et al., 2020; Mansour et al., 2022). The LPNEB and LEDS showed strong genomic correlations with 2 major ketone bodies, LB_BHB and LACE, which

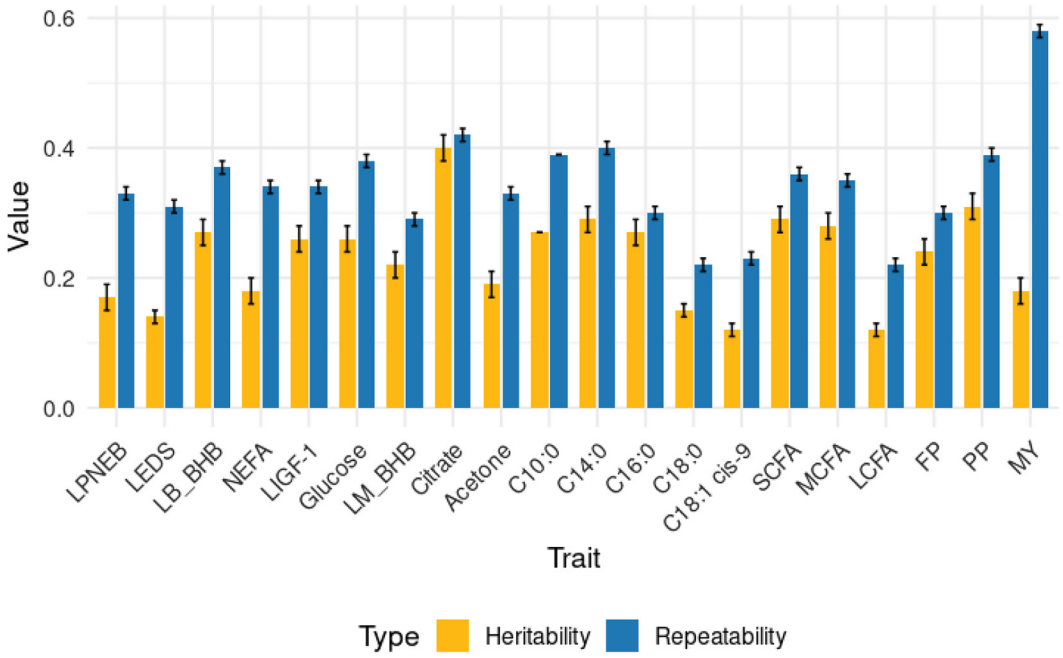

Figure 1. Heritability and repeatability of 20 traits estimated based on 20 univariate repeatability models. LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHB = $\log_{10}$-transformed blood β-hydroxybutyrate predicted by MIR spectra; NEFA = blood nonesterified fatty acids predicted by MIR spectra; LIGF-1 = $\log_{10}$-transformed blood IGF-1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHB = $\log_{10}$-transformed milk β-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; LACE = $\log_{10}$-transformed milk acetone predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; C14:0 = milk myristic acid predicted by MIR spectra; C16:0 = milk palmitic acid predicted by MIR spectra; C18:0 = milk stearic acid predicted by MIR spectra; C18:1 *cis*-9 = milk oleic acid predicted by MIR spectra; SCFA = milk short-chain fatty acids predicted by MIR spectra; MCFA = milk medium-chain fatty acids predicted by MIR spectra; LCFA = milk long-chain fatty acids predicted by MIR spectra; FP = milk fat percentage predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

are synthesized in the liver from NEFA during NEB. Our findings are in close agreement with the metabolic pathway, whereby elevated NEFA levels stimulate the hepatic production of BHB and acetone (Guliński, 2021; Lisuzzo et al., 2022; Martens, 2023). The C18:1 *cis*-9, a monounsaturated long-chain FA, showed strong positive correlations with both LPNEB and LEDS. Increased fat mobilization during NEB results in higher levels of LCFA, including C18:1 *cis*-9, due to enhanced hepatic β-oxidation (Bastin et al., 2012; Churakov et al., 2021), which supports our findings.

Correlation Between LPNEB and 19 Traits Across Different Chromosomes

The genomic correlations estimated between LPNEB and 19 traits across different chromosomal regions are presented in Figure 3. The significance tests for the genomic correlations shown in Figure 3 are presented in Supplemental Table S1, file 2 (see Notes). At the All_BTA, LPNEB showed strong positive genetic correlations with LEDS (0.81), LB_BHB (0.55), and NEFA (0.80). These observations align well with previous studies that have reported elevated lipid mobilization during periods of NEB in early lactation (Ospina et al., 2010; Andjelić et al., 2022).

Through chromosome-specific analyses, BTA25, which showed the highest MeanAbsCorr (0.51), was identified as a key chromosome associated with LPNEB. The chromosome-specific MeanAbsCorr values for LPNEB are provided in Supplemental Table S2, file 3 (see Notes). The LPNEB on BTA25 showed the strongest genetic correlations with LEDS (0.88), LB_BHB (0.74), and NEFA (0.87). Ha et al. (2015) found that genes related to BHB and NEFA were involved in key energy metabolism pathways. Some of these genes (e.g., *WBSCR22*) are located on BTA25, supporting the role of this chromosome in energy balance during early lactation. In the present study, LPNEB also exhibited strong positive genetic correlations with C18:1 *cis*-9 (0.75) and LCFA (0.62) on BTA25, supported by the known effect of NEB during lactation on milk FA composition. Previous studies have reported associations between energy balance during early lactation in dairy cows and genomic regions

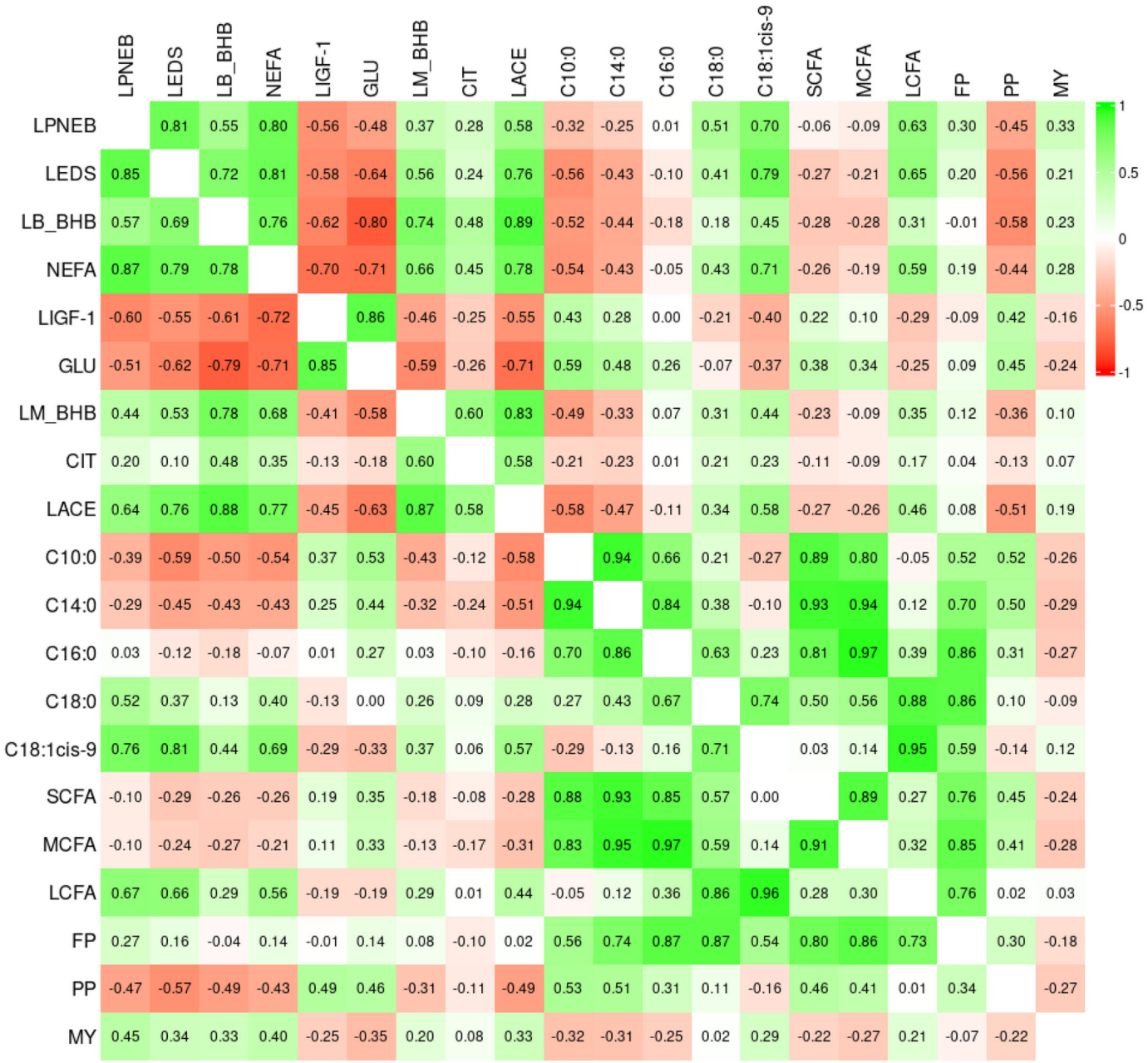

Figure 2. Genomic correlations among 20 traits. Values in the upper diagonal were calculated based on SNP effects from 20 univariate repeatability models, while values in the lower diagonal were derived from 20-traits repeatability model. The SE for the upper diagonal elements of the correlation matrix ranged from 0.0001 to 0.0013. The 20 traits include: LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHB = $\log_{10}$-transformed blood β-hydroxybutyrate predicted by MIR spectra; NEFA = blood nonesterified fatty acids predicted by MIR spectra; LIGF-1 = $\log_{10}$-transformed blood IGF-1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHB = $\log_{10}$-transformed milk β-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; LACE = $\log_{10}$-transformed milk acetone predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; C14:0 = milk myristic acid predicted by MIR spectra; C16:0 = milk palmitic acid predicted by MIR spectra; C18:0 = milk stearic acid predicted by MIR spectra; C18:1 *cis*-9 = milk oleic acid predicted by MIR spectra; SCFA = milk short-chain fatty acids predicted by MIR spectra; MCFA = milk medium-chain fatty acids predicted by MIR spectra; LCFA = milk long-chain fatty acids predicted by MIR spectra; FP = milk fat percentage predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

located on BTA1, BTA15, and BTA16 (Tetens et al., 2013; Krattenmacher et al., 2019). The identification of BTA25 as potential regulatory regions represents novel findings, indicating the possible involvement of previously unrecognized genetic mechanisms underlying the regulation of the NEB.

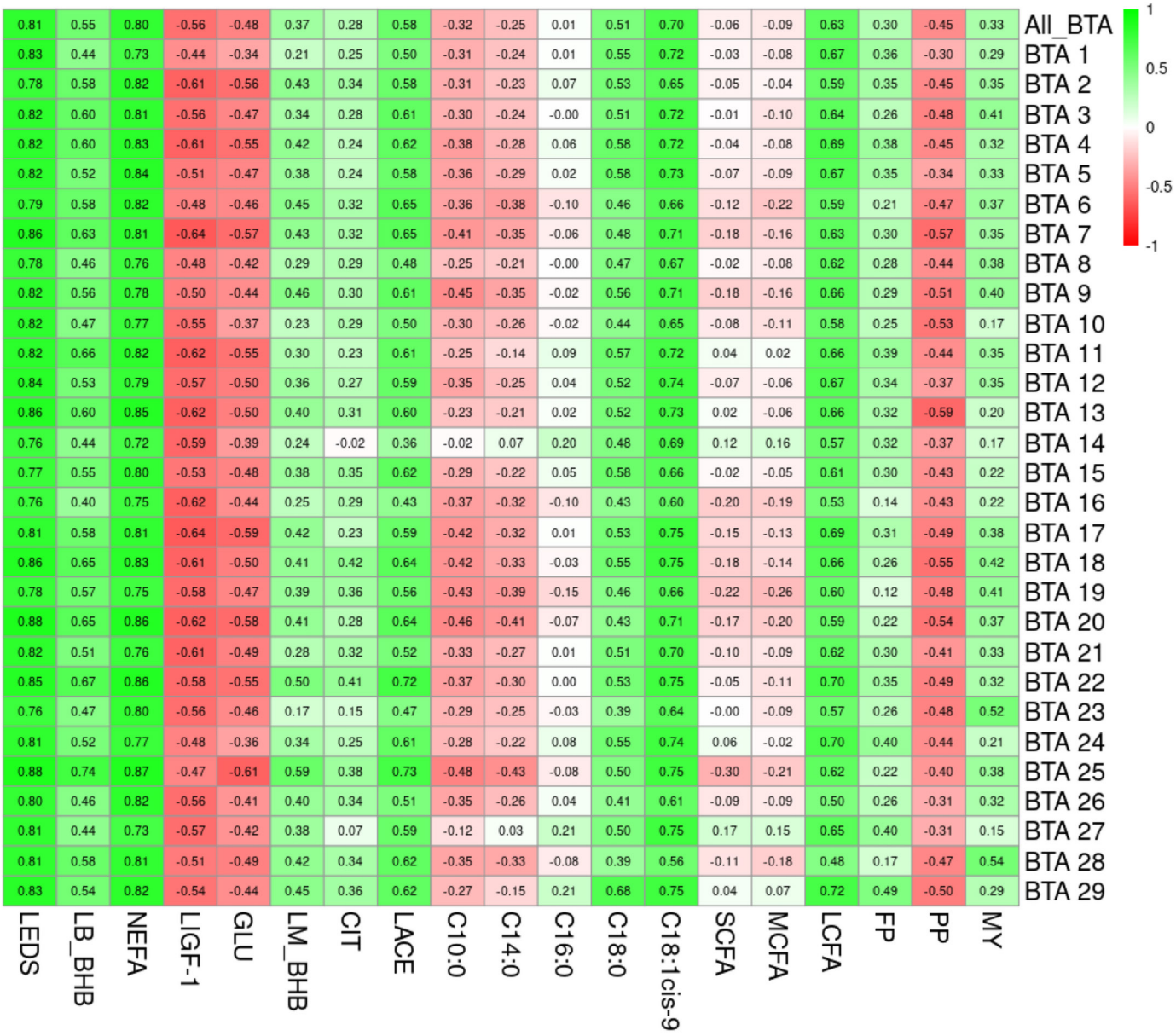

Figure 3. Genomic correlations between logarithm probability negative energy balance (LPNEB) and 19 other traits based on SNP effects. The first row represents genomic correlations estimated using SNP effects across all 29 chromosomes. Rows 2 to 30 represent genomic correlations estimated using SNP effects from each individual chromosome. The SE of the correlation matrix ranged from 0.0003 to 0.0013. The 20 traits include: LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHB = $\log_{10}$-transformed blood β-hydroxybutyrate predicted by MIR spectra; NEFA = blood nonesterified fatty acids predicted by MIR spectra; LIGF-1 = $\log_{10}$-transformed blood IGF-1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHB = $\log_{10}$-transformed milk β-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; LACE = $\log_{10}$-transformed milk acetone predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; C14:0 = milk myristic acid predicted by MIR spectra; C16:0 = milk palmitic acid predicted by MIR spectra; C18:0 = milk stearic acid predicted by MIR spectra; C18:1 *cis*-9 = milk oleic acid predicted by MIR spectra; SCFA = milk short-chain fatty acids predicted by MIR spectra; MCFA = milk medium-chain fatty acids predicted by MIR spectra; LCFA = milk long-chain fatty acids predicted by MIR spectra; FP = milk fat percentage predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

Correlation Between LEDS and 19 Traits Across Different Chromosomes

Figure 4 presents the genomic correlations between LEDS and 19 traits across different chromosomal regions. All significance tests of the genomic correlations presented in Figure 4 are reported in Supplemental Table S3, file 4 (see Notes). At the All_BTA, LEDS showed

strong positive genetic correlations with LPNEB (0.81), NEFA (0.81), and C18:1 *cis*-9 (0.79).

The BTA19, which exhibited the highest MeanAbsCorr (0.58), was identified as the key chromosome associated with LEDS. The BTA25 showed the second highest MeanAbsCorr (0.57). The chromosome-specific MeanAbsCorr values for LEDS are provided in Supplemental Table S5, file 5 (see Notes). Some SNPs on BTA19 associated with energy balance during early lactation in dairy cows were identified by Tetens et al. (2013). The LEDS on BTA19 showed the strongest genetic correlations with C18:1 *cis*-9 (0.84), LACE (0.84), and NEFA (0.82). Knutsen et al. (2022) reported that BTA19 is significantly associated with C18:1 *cis*-9. These findings suggest that BTA19 may play a central role in the regulation of NEB in dairy cows during early lactation.

Independent Contributions of 8 Selected Traits to LPNEB and LEDS

As shown in Figure 5, the selected 8 traits (NEFA, IGF-1, LB_BHB, LM_BHB, CIT, C10:0, PP, and MY) showed both similarities and differences in their independent contributions to LPNEB and LEDS. Results were revealed to be different and somewhat complementary to observed correlations also reported in this study.

The NEFA exhibited strong genomic correlations with both LPNEB and LEDS and also accounted for a major proportion in independent contribution analysis. These results suggest that NEFA is not only highly genomically correlated with LPNEB and LEDS but may also exert direct genetic effects on both traits. As a key biomarker of fat mobilization during NEB, the underlying biological role of NEFA aligns with its strong genomic correlations and substantial independent contributions observed in this study, further supporting its central function in the regulation of LPNEB (Mehtiö et al., 2020; Mansour et al., 2022).

Although BHB (including LB_BHB and LM_BHB) showed moderate positive genomic correlations with both LPNEB and LEDS, its independent contributions were generally low. This may be due to collinearity with stronger traits such as NEFA. As a ketone body derived from FA mobilization, BHB likely reflects metabolic outcomes rather than acting as a direct genetic driver of NEB. The IGF-1 showed moderate genomic correlations with NEB, their low independent contributions indicate that their influence may be indirect. Results for PP showed consistent negative genomic correlations and independent contributions for both LPNEB and LEDS, implying a potential opposite genetic relationship with NEB. Other traits, such as C10:0, CIT, and MY, displayed weak correlations or small independent contributions, suggesting these traits appear to have limited direct genetic effects on NEB and may play a secondary or supporting role.

In summary, NEFA showed the strongest genomic correlation and independent contribution with both LPNEB and LEDS, indicating that LPNEB and LEDS may both serve as good proxy traits for NEB. However, the current prediction models for LPNEB and NEFA show limited accuracy, which restricts their applicability in breeding programs.

Compared with single biomarkers such as NEFA, LEDS emerged as a more robust proxy by combining multiple MIR-predicted traits, achieving higher predictive accuracy (0.99) and sufficient h^2. The LEDS can be accurately predicted from routinely collected MIR spectra in DHI centers, enabling large-scale monthly recording similar to other milk traits such as FP. Therefore, we recommend its implementation as a proxy trait for NEB in dairy cattle breeding programs, although further work is required to evaluate LEDS genetic relationships with existing selection traits and its effect on overall breeding objectives.

There are still some aspects of this study that need to be improved. First, except for MY, all phenotypes were predicted from MIR spectra rather than directly measured. Because these predicted traits were derived from the same spectra, they may share prediction errors, which could artificially inflate the estimated genomic correlations. This potential bias should be acknowledged when interpreting the results; however, the overall correlation patterns were consistent with known genetic relationships, supporting the robustness of the main findings. We recommend that future studies include external validation with large datasets of measured phenotypes wherever possible to minimize this bias. Second, the potential functional relevance of BTA19 and BTA25, which were identified as key chromosomes associated with NEB and related traits in this study, has not yet been explored through functional annotation or experimental analysis. In addition, it should be noted that this study focused on identifying key chromosomes with shared effects on NEB and its 19 related traits. Third, genomic correlations were estimated using correlations of SNP effects from single-trait analyses, which do not account for sampling covariances among traits. Although multitrait models with individual-level data, such as multitrait ssGREML, provide more accurate estimates, their application remains computationally challenging for large numbers of traits. Further methodological advances are needed to address this limitation.

CONCLUSIONS

This study compared the genetic architecture of LPNEB and LEDS and their relationships with 15 biomarkers

Figure 4. Genomic correlations between logarithm probability energy deficiency score (LEDS) and 19 other traits based on SNP effects. The first row represents genomic correlations estimated using SNP effects across all 29 chromosomes, Rows 2 to 30 represent genomic correlations estimated using SNP effects from individual chromosomes. The SE of the correlation matrix ranged from 0.0003 to 0.0013. The 20 traits include: LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHB = $\log_{10}$-transformed blood β-hydroxybutyrate predicted by MIR spectra; NEFA = blood nonesterified fatty acids predicted by MIR spectra; LIGF-1 = $\log_{10}$-transformed blood IGF-1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHB = $\log_{10}$-transformed milk β-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; LACE = $\log_{10}$-transformed milk acetone predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; C14:0 = milk myristic acid predicted by MIR spectra; C16:0 = milk palmitic acid predicted by MIR spectra; C18:0 = milk stearic acid predicted by MIR spectra; C18:1 *cis*-9 = milk oleic acid predicted by MIR spectra; SCFA = milk short-chain fatty acids predicted by MIR spectra; MCFA = milk medium-chain fatty acids predicted by MIR spectra; LCFA = milk long-chain fatty acids predicted by MIR spectra; FP = milk fat percentage predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

and 3 production traits through genomic correlation and independent contribution analyses. Our results show SNP effects estimated from single-trait models can be used to quickly calculate genomic correlations for 20 traits. Among all traits, NEFA showed strong genomic correlations to both LPNEB and LEDS, indicating potential roles as direct genetic drivers of NEB. Chromosome-level analyses identified key chromosomes such as BTA19 and

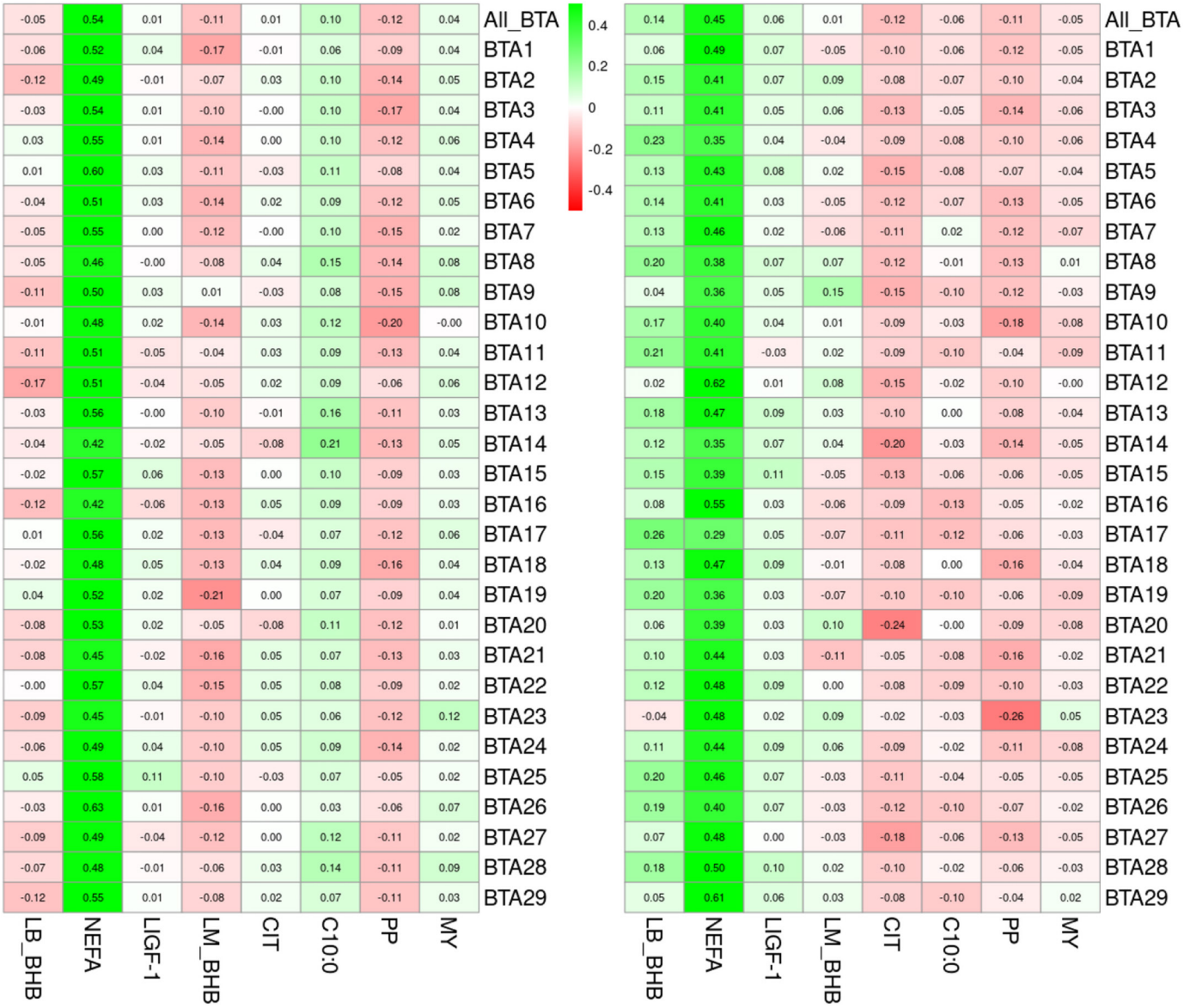

Figure 5. Independent contribution of 8 traits to LPNEB (left) and LEDS (right). The first row displays the independent contributions across all 29 chromosomes. Rows 2 to 30 represent independent contributions estimated from each individual chromosome. The 8 traits include: LB_BHB = $\log_{10}$-transformed blood β-hydroxybutyrate predicted by MIR spectra; NEFA = blood nonesterified fatty acids predicted by MIR spectra; LIGF-1 = $\log_{10}$-transformed blood IGF-1 predicted by MIR spectra; LM_BHB = $\log_{10}$-transformed milk β-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

BTA25. Independent contribution analysis further indicated that NEFA contributed most strongly to LPNEB and LEDS. Overall, integrating genomic correlation and contribution analyses provides a clearer understanding of the genetic basis of NEB and its related traits in dairy cows.

NOTES

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Hongqing Hu (no. 202207650049). The computation resources of the University of Liège-Gembloux Agro-Bio Tech (ULiège-GxABT, Gembloux, Belgium) were partly supported by the Fonds de la Recherche Scientifique (FRS-FNRS, Brussels, Belgium), which also provided support through the PDR projects "HTwoTHI" (grant number T.W005.23)

and "DEEPSELECT" (grant number T.0095.19). Author contributions are as follows: Hongqing Hu, funding acquisition, formal analysis, and writing (original draft); Sébastien Franceschini, Clément Grelet, and Yansen Chen, data curation and writing (review and editing); Pauline Lemal, Hadi Atashi, Katrien Wijnrocx, and Soyeurt Hélène, validation and writing (review and editing); and Nicolas Gengler, conceptualization, funding acquisition, methodology, project administration, software, supervision, validation, and writing (review and editing). The supplemental materials for this article are available at https://github.com/Hongqing92/LPNEB. No human or animal subjects were used, so this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: C10:0 = milk decanoic acid; C14:0 = milk myristic acid; C16:0 = milk palmitic acid; C18:0 = milk stearic acid; CIT = milk citrate; EDS = energy deficiency score; FA = fatty acids; FDR = false discovery rate; FP = milk fat percentage; GLU = blood glucose; LACE = $\log_{10}$-transformed milk acetone; LB_BHB = $\log_{10}$-transformed blood BHB; LCFA = milk long-chain FA; LEDS = logit-transformed novel energy deficiency score; LIGF-1 = $\log_{10}$-transformed blood IGF-1; LM_BHB = $\log_{10}$-transformed milk BHB; LPNEB = logit-transformed predicted NEB; MCFA = milk medium-chain FA; MeanAbsCorr = mean absolute genetic correlation; MIR = mid-infrared; MY = milk yield; NEB = negative energy balance; NEFA = nonesterified fatty acids; PEB = predicted energy balance; PP = milk protein percentage; SCFA = milk short-chain FA; VIF = variance inflation factor.

REFERENCES

Aernouts, B., I. Adriaens, J. Diaz-Olivares, W. Saeys, P. Mäntysaari, T. Kokkonen, T. Mehtiö, S. Kajava, P. Lidauer, M. H. Lidauer, and M. Pastell. 2020. Mid-infrared spectroscopic analysis of raw milk to predict the blood nonesterified fatty acid concentrations in dairy cows. J. Dairy Sci. 103:6422–6438. https://doi.org/10.3168/jds.2019-17952.

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

Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. Proceedings of the 10th world congress of genetics applied to livestock production, Vancouver, BC, Canada. Accessed Apr. 10, 2025. http://www.ainfo.inia.uy/digital/bitstream/item/15445/1/Aguilar-et-al.-2014.-WCGALP.pdf.

Andjelić, B., R. Djoković, M. Cincović, S. Bogosavljević-Bošković, M. Petrović, J. Mladenović, and A. Čukić. 2022. Relationships between milk and blood biochemical parameters and metabolic status in dairy cows during lactation. Metabolites 12:733. https://doi.org/10.3390/metabo12080733.

Atashi, H., Y. Chen, S. Vanderick, X. Hubin, and N. Gengler. 2024. Single-step genome-wide association analyses for milk urea concentration in Walloon Holstein cows. J. Dairy Sci. 107:3020–3031. https://doi.org/10.3168/jds.2023-23902.

Atashi, H., Y. Chen, H. Wilmot, S. Vanderick, X. Hubin, H. Soyeurt, and N. Gengler. 2023. Single-step genome-wide association for selected milk fatty acids in Dual-Purpose Belgian Blue cows. J. Dairy Sci. 106:6299–6315. https://doi.org/10.3168/jds.2022-22432.

Bastin, C., D. P. Berry, H. Soyeurt, and N. Gengler. 2012. Genetic correlations of days open with production traits and contents in milk of major fatty acids predicted by mid-infrared spectrometry. J. Dairy Sci. 95:6113–6121. https://doi.org/10.3168/jds.2012-5361.

Bastin, C., N. Gengler, and H. Soyeurt. 2011. Phenotypic and genetic variability of production traits and milk fatty acid contents across days in milk for Walloon Holstein first-parity cows. J. Dairy Sci. 94:4152–4163. https://doi.org/10.3168/jds.2010-4108.

Becker, V. A. E., E. Stamer, H. Spiekers, and G. Thaller. 2021. Residual energy intake, energy balance, and liability to diseases: Genetic parameters and relationships in German Holstein dairy cows. J. Dairy Sci. 104:10970–10978. https://doi.org/10.3168/jds.2021-20382.

Benedet, A., A. Costa, M. De Marchi, and M. Penasa. 2020. Heritability estimates of predicted blood β-hydroxybutyrate and nonesterified fatty acids and relationships with milk traits in early-lactation Holstein cows. J. Dairy Sci. 103:6354–6363. https://doi.org/10.3168/jds.2019-17916.

Billa, P. A., Y. Faulconnier, T. Larsen, C. Leroux, and J. A. A. Pires. 2020. Milk metabolites as noninvasive indicators of nutritional status of mid-lactation Holstein and Montbéliarde cows. J. Dairy Sci. 103:3133–3146. https://doi.org/10.3168/jds.2019-17466.

Buttchereit, N., E. Stamer, W. Junge, and G. Thaller. 2011. Short communication: Genetic relationships among daily energy balance, feed intake, body condition score, and fat to protein ratio of milk in dairy cows. J. Dairy Sci. 94:1586–1591. https://doi.org/10.3168/jds.2010-3396.

Chen, Y., H. Hu, H. Atashi, C. Grelet, K. Wijnrocx, P. Lemal, and N. Gengler. 2024. Genetic analysis of milk citrate predicted by milk mid-infrared spectra of Holstein cows in early lactation. J. Dairy Sci. 107:3047–3061. https://doi.org/10.3168/jds.2023-23903.

Churakov, M., J. Karlsson, A. Edvardsson Rasmussen, and K. Holtenius. 2021. Milk fatty acids as indicators of negative energy balance of dairy cows in early lactation. Animal 15:100253. https://doi.org/10.1016/j.animal.2021.100253.

Esposito, G., P. C. Irons, E. C. Webb, and A. Chapwanya. 2014. Interactions between negative energy balance, metabolic diseases, uterine health, and immune response in transition dairy cows. Anim. Reprod. Sci. 144:60–71. https://doi.org/10.1016/j.anireprosci.2013.11.007.

Franceschini, S., N. Gengler, and H. Soyeurt. 2024. Combining high throughput phenotypes to study complex traits: A case-study of negative energy balance using milk mid-infrared based predictions. The 1st meeting European Network on Livestock Phenomics. https://hdl.handle.net/2268/329610.

Franceschini, S., C. Grelet, J. Leblois, N. Gengler, and H. Soyeurt. 2022. Can unsupervised learning methods applied to milk recording big data provide new insights into dairy cow health? J. Dairy Sci. 105:6760–6772. https://doi.org/10.3168/jds.2022-21975.

Gardner, K. M., and R. G. Latta. 2007. Shared quantitative trait loci underlying the genetic correlation between continuous traits. Mol. Ecol. 16:4195–4209. https://doi.org/10.1111/j.1365-294X.2007.03499.x.

Gohary, K., M. W. Overton, M. von Massow, S. J. LeBlanc, K. D. Lissemore, and T. F. Duffield. 2016. The cost of a case of subclinical ketosis in Canadian dairy herds. Can. Vet. J. 57:728–732.

Grelet, C., J. A. Fernández Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of milk mid-infrared spectra from a European dairy network. J. Dairy Sci. 98:2150–2160. https://doi.org/10.3168/jds.2014-8764.

Grelet, C., A. Vanlierde, F. Dehareng, E. Froidmont, and E. Gplus. Consortium. 2017, Prediction of energy status of dairy cows using MIR milk spectra. Page 403 in Book of Abstracts of the 68th Annual

Meeting of the European Federation of Animal Science, Tallin, Wageningen Academic Publishers. https://hdl.handle.net/2268/224000.

Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. L. Ingvartsen, M. Crowe, M. T. Sorensen, E. Froidmont, C. P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, and E. Gplus. Consortium, and F. Dehareng. 2019. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. Animal 13:649–658. https://doi.org/10.1017/S1751731118001751.

Guliński, P. 2021. Ketone bodies – causes and effects of their increased presence in cows' body fluids: A review. Vet. World 14:1492–1503. https://doi.org/10.14202/vetworld.2021.1492-1503.

Ha, N.-T., J. J. Gross, A. van Dorland, J. Tetens, G. Thaller, M. Schlather, R. Bruckmaier, and H. Simianer. 2015. Gene-based mapping and pathway analysis of metabolic traits in dairy cows. PLoS One 10:e0122325. https://doi.org/10.1371/journal.pone.0122325.

Hammon, D. S., I. M. Evjen, T. R. Dhiman, J. P. Goff, and J. L. Walters. 2006. Neutrophil function and energy status in Holstein cows with uterine health disorders. Vet. Immunol. Immunopathol. 113:21–29. https://doi.org/10.1016/j.vetimm.2006.03.022.

Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: A review with focus on the situation in the Nordic countries. Livest. Prod. Sci. 64:95–106. https://doi.org/10.1016/S0301-6226(99)00128-1.

Ho, P. N., L. C. Marett, W. J. Wales, M. Axford, E. M. Oakes, and J. E. Pryce. 2019. Predicting milk fatty acids and energy balance of dairy cows in Australia using milk mid-infrared spectroscopy. Anim. Prod. Sci. 60:164–168. https://doi.org/10.1071/AN18532.

Hu, H., S. Franceschini, P. Lemal, C. Grelet, Y. Chen, H. Atashi, K. Wijnrocx, H. Soyeurt, and N. Gengler. 2025. Exploring the relationship between predicted negative energy balance and its biomarkers of Holstein cows in first-parity early lactation. J. Dairy Sci. 108:5433–5447. https://doi.org/10.3168/jds.2024-25932.

ICAR (International Committee for Animal Recording). 2022. Procedure 2 of Section 2 of ICAR Guidelines - Computing of Accumulated Lactation Yield. Accessed Dec. 29, 2023. https://www.icar.org/Guidelines/02-Procedure-2-Computing-Lactation-Yield.pdf.

Knob, D. A., A. Thaler Neto, H. Schweizer, A. C. Weigand, R. Kappes, and A. M. Scholz. 2021. Energy balance indicators during the transition period and early lactation of purebred Holstein and Simmental cows and their crosses. Animals (Basel) 11:309. https://doi.org/10.3390/ani11020309.

Knutsen, T. M., H. G. Olsen, I. A. Ketto, K. K. Sundsaasen, A. Kohler, V. Tafintseva, M. Svendsen, M. P. Kent, and S. Lien. 2022. Genetic variants associated with two major bovine milk fatty acids offer opportunities to breed for altered milk fat composition. Genet. Sel. Evol. 54:35. https://doi.org/10.1186/s12711-022-00731-9.

Koeck, A., J. Jamrozik, F. S. Schenkel, R. K. Moore, D. M. Lefebvre, D. F. Kelton, and F. Miglior. 2014. Genetic analysis of milk β-hydroxybutyrate and its association with fat-to-protein ratio, body condition score, clinical ketosis, and displaced abomasum in early first lactation of Canadian Holsteins. J. Dairy Sci. 97:7286–7292. https://doi.org/10.3168/jds.2014-8405.

Krattenmacher, N., G. Thaller, and J. Tetens. 2019. Analysis of the genetic architecture of energy balance and its major determinants dry matter intake and energy-corrected milk yield in primiparous Holstein cows. J. Dairy Sci. 102:3241–3253. https://doi.org/10.3168/jds.2018-15480.

Lisuzzo, A., L. Laghi, V. Faillace, C. Zhu, B. Contiero, M. Morgante, E. Mazzotta, M. Gianesella, and E. Fiore. 2022. Differences in the serum metabolome profile of dairy cows according to the BHB concentration revealed by proton nuclear magnetic resonance spectroscopy (^{1}H-NMR). Sci. Rep. 12:2525. https://doi.org/10.1038/s41598-022-06507-x.

Macrae, A. I., E. Burrough, J. Forrest, A. Corbishley, G. Russell, and D. J. Shaw. 2019. Prevalence of excessive negative energy balance in commercial United Kingdom dairy herds. Vet. J. 248:51–57. https://doi.org/10.1016/j.tvjl.2019.04.001.

Mansour, U. M., H. E. Belal, and R. M. Dohreig. 2022. Biomarkers for negative energy balance and fertility in early lactating dairy cows. Ger. J. Vet. Res. 2:11–16. https://doi.org/10.51585/gjvr.2022.2.0031.

Martens, H. 2023. Invited review: Increasing milk yield and negative energy balance: A Gordian knot for dairy cows? Animals (Basel) 13:3097. https://doi.org/10.3390/ani13193097.

Mehtiö, T., P. Mäntysaari, E. Negussie, A. M. Leino, J. Pösö, E. A. Mäntysaari, and M. H. Lidauer. 2020. Genetic correlations between energy status indicator traits and female fertility in primiparous Nordic Red Dairy cattle. Animal 14:1588–1597. https://doi.org/10.1017/S1751731120000439.

Mekuriaw, Y. 2023. Negative energy balance and its implication on productive and reproductive performance of early lactating dairy cows: Review paper. J. Appl. Anim. Res. 51:220–228. https://doi.org/10.1080/09712119.2023.2176859.

Miglior, F., A. Fleming, F. Malchiodi, L. F. Brito, P. Martin, and C. F. Baes. 2017. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. J. Dairy Sci. 100:10251–10271. https://doi.org/10.3168/jds.2017-12968.

Miles, J. 2005. Tolerance and variance inflation factor. Encyclopedia of Statistics in Behavioral Science. Wiley. https://doi.org/10.1002/0470013192.bsa683.

Misztal, I., S. Tsuruta, D. A. L. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. Manual for BLUPF90 family programs. University of Georgia. Accessed Apr. 9, 2025. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Oikonomou, G., G. E. Valergakis, G. Arsenos, N. Roubies, and G. Banos. 2008. Genetic profile of body energy and blood metabolic traits across lactation in primiparous Holstein cows. J. Dairy Sci. 91:2814–2822. https://doi.org/10.3168/jds.2007-0965.

Ospina, P. A., D. V. Nydam, T. Stokol, and T. R. Overton. 2010. Association between the proportion of sampled transition cows with increased nonesterified fatty acids and β-hydroxybutyrate and disease incidence, pregnancy rate, and milk production at the herd level. J. Dairy Sci. 93:3595–3601. https://doi.org/10.3168/jds.2010-3074.

Paiva, J. T., R. R. Mota, P. S. Lopes, H. Hammami, S. Vanderick, H. R. Oliveira, R. Veroneze, F. Fonseca e Silva, and N. Gengler. 2022. Random regression test-day models to describe milk production and fatty acid traits in first lactation Walloon Holstein cows. J. Anim. Breed. Genet. 139:398–413. https://doi.org/10.1111/jbg.12673.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478.

Smith, S. L., S. J. Denholm, M. P. Coffey, and E. Wall. 2019. Energy profiling of dairy cows from routine milk mid-infrared analysis. J. Dairy Sci. 102:11169–11179. https://doi.org/10.3168/jds.2018-16112.

Spurlock, D. M., J. C. M. Dekkers, R. Fernando, D. A. Koltes, and A. Wolc. 2012. Genetic parameters for energy balance, feed efficiency, and related traits in Holstein cattle. J. Dairy Sci. 95:5393–5402. https://doi.org/10.3168/jds.2012-5407.

Tetens, J., T. Seidenspinner, N. Buttchereit, and G. Thaller. 2013. Whole-genome association study for energy balance and fat/protein ratio in German Holstein bull dams. Anim. Genet. 44:1–8. https://doi.org/10.1111/j.1365-2052.2012.02357.x.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. 94:73–83. https://doi.org/10.1017/S0016672312000274.

Wathes, D. C., M. Fenwick, Z. Cheng, N. Bourne, S. Llewellyn, D. G. Morris, D. Kenny, J. Murphy, and R. Fitzpatrick. 2007. Influence of negative energy balance on cyclicity and fertility in the high producing dairy cow. Theriogenology 68:S232–S241. https://doi.org/10.1016/j.theriogenology.2007.04.006.

Whitfield, R. G., M. E. Gerger, and R. L. Sharp. 1987. Near-infrared spectrum qualification via Mahalanobis distance determination. Appl. Spectrosc. 41:1204–1213. https://doi.org/10.1366/0003702874447572.

Wiggans, G. R., T. S. Sonstegard, P. M. VanRaden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, F. S. Schenkel, and C. P. van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of

genotypes used in genomic evaluation of dairy cattle in the United States and Canada. J. Dairy Sci. 92:3431–3436. https://doi.org/10.3168/jds.2008-1758.