

# Using single-step genome-wide association analyses to compare predicted negative energy balance and a novel energy deficiency score in early-lactation Holstein cows

Hongqing Hu,<sup>1</sup> Hadi Atashi,<sup>1,2</sup> S. Franceschini,<sup>1</sup> Pauline Lemal,<sup>1</sup> Clément Grelet,<sup>3</sup> Yansen Chen,<sup>1</sup> Katrien Wijnrocx,<sup>1</sup> Hélène Soyeurt,<sup>1</sup> and Nicolas Gengler<sup>1\*</sup>

**Abstract:** This study aimed to compare the genetic architectures of logit-transformed predicted negative energy balance (LPNEB) and a novel energy deficiency score (logit-transformed energy deficiency score [LEDS]) as 2 mid-infrared–derived proxies of negative energy balance in early-lactation dairy cows. A total of 30,634 records from 25,287 first-parity Holstein cows across 508 herds distributed in Walloon region of Belgium were analyzed. Genotypic data of 566,170 SNPs were available for 3,757 animals. Single-step GWAS, combined with a 50-SNP sliding window approach, was employed to explore the genetic architectures of LPNEB and LEDS. The top 10 genomic regions for LPNEB and LEDS were identified across multiple chromosomes, with 3 shared regions (BTA 1, 5, and 16). Despite these overlaps, each trait exhibited unique loci, supporting distinct genetic architectures. Positional candidate gene analyses identified 17 genes for LPNEB and 10 for LEDS, with 6 being in common. Gene Ontology enrichment analyses were then performed to explore their biological functions, although LPNEB was primarily associated with energy metabolism regulation and metabolic adaptation, whereas LEDS integrated neuronal signaling into energy homeostasis. The QTL enrichment highlighted significant associations with fertility and SCS, reinforcing a genetic basis for energy balance. These findings improve our understanding of the genetic background of LPNEB and LEDS, thereby providing new insights into the mechanisms underlying energy balance in dairy cattle.

Negative energy balance (NEB) is a well-recognized metabolic challenge in early-lactation dairy cows, occurring when energy intake is insufficient to meet the high demands of maintenance and milk production (Churakov et al., 2021). Prolonged NEB is associated with metabolic disorders such as ketosis and fatty liver (Esposito et al., 2014) and is linked to reduced reproductive performance, milk yield, and immune functions (Hammon et al., 2006; Pérez-Báez et al., 2019; Wathes et al., 2007). These consequences not only have an impact on cow health and welfare, but also impose substantial economic costs on dairy farms (Gohary et al., 2016; Miglior et al., 2017). Therefore, effective monitoring of NEB is crucial for maintaining herd health and productivity. Direct measurement of NEB on a large scale is challenging due to its high cost (Zachut et al., 2020). Mid-infrared (MIR) spectral can be considered as an efficient approach to predict NEB status in dairy cows; however, the coefficients of determination of the different prediction models vary across studies (from 0.48 to 0.78; Grelet et al., 2017; Smith et al., 2019).

A promising alternative is the energy deficiency score (EDS), derived from MIR data, which provides a noninvasive and cost-effective method for assessing energy status (Franceschini et al., 2020). The EDS integrates multiple metabolic markers, making it a more reliable indicator of NEB than individual biomarkers. Cows in the EDS group show lower BWt and DMI, along with higher

BHB and acetone in milk (Franceschini et al., 2022). These changes suggest increased fat mobilization, a key sign of NEB induced metabolic stress. The genetic correlations with blood biomarkers further support this finding, with logit-transformed EDS (LEDS) being strongly positively correlated with logit-transformed predicted NEB (LPNEB; 0.85), blood BHB (0.69), and blood non-esterified fatty acid (0.79), while showing a negative correlation with blood glucose (−0.62; Hu et al., 2025). Although both LPNEB and LEDS are MIR-based proxies for the NEB, LEDS may offer a more integrative and scalable alternative to LPNEB for genetic and management applications. Our recent study suggests moderate heritability for LPNEB and LEDS (Hu et al., 2025), highlighting the potential for genetic selection. Identifying genomic regions associated with LPNEB and LEDS could improve understanding of the relationship of these traits to metabolic adaptation and support their potential use in breeding strategies for better energy efficiency. Single-step GWAS (ssGWAS) provide a powerful approach to identify genomic regions associated with complex metabolic traits by integrating genotype and phenotype data from large populations.

To compare the genetic architectures of LPNEB and LEDS in early-lactation Holstein cows, this study (1) performed ssGWAS to identify genomic regions associated with LPNEB and LEDS,

<sup>1</sup>TERRA Teaching and Research Center, University of Liège, Gembloux Agro-Bio Tech (ULiège-GxABT), 5030 Gembloux, Belgium, <sup>2</sup>Department of Animal Science, Shiraz University, 71441-13131 Shiraz, Iran, <sup>3</sup>Walloon Agricultural Research Center (CRA-W), 5030 Gembloux, Belgium. \*Corresponding author: [nicolas.gengler@uliege.be](mailto:nicolas.gengler@uliege.be). © 2025, The Authors. Published by Elsevier Inc. on behalf of the American Dairy Science Association®. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>). Received April 04, 2025. Accepted July 06, 2025.

The list of standard abbreviations for JDSC is available at [adsa.org/jdsc-abbreviations-25](https://adsa.org/jdsc-abbreviations-25). Nonstandard abbreviations are available in the Notes.

and (2) conducted comparative functional annotation of selected genomic regions.

The data used in this study were explained previously by Hu et al. (2025). Briefly, the dataset consisted of 30,634 records on 25,287 first-parity Holstein cows across 508 herds in Walloon region of Belgium. Some cows had multiple test-day records because MIR spectra and predicted traits were collected on different days during the early lactation period (DIM 5 to 50). To focus on early lactation, records were restricted to the period of 5 to 50 DIM. The MIR-based energy balance prediction equations used for trait estimation were developed and validated in previous studies (Grelet et al., 2017). The EDS was defined using an agglomerative hierarchical clustering approach based on 27 MIR-derived predictors (Franceschini et al., 2022). For novel observations, EDS was directly predicted from MIR spectral data, ensuring consistency with previously established methodologies. The 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 used pedigree included 76,340 animals, among which 3,757 had genotypic data for 566,170 SNPs. The genotype data were obtained from a subset of the dataset described in Chen et al. (2024). The LPNEB and LEDS formulas are as follows:

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

where

$$\text{PNEB} = 1 - \frac{\text{PEB} - \text{PEB}_{\text{minimum}}}{\text{PEB}_{\text{maximum}} - \text{PEB}_{\text{minimum}}}, \text{ and}$$

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

with PEB being the energy balance predicted by mid-infrared spectra and PNEB being the NEB predicted by mid-infrared spectra. Similar to Hu et al. (2025), a univariate repeatability model was used to estimate variance components and GEBV for LPNEB and LEDS in first-parity cows. The model included fixed effects of herd  $\times$  calving year, calving age (10 classes), calving month (12 classes), and DIM classes (4 levels). In addition, standardized DIM (centered and scaled to mean 0 and SD 1) and its quadratic term were included to model linear and nonlinear effects of DIM during early lactation. Random effects were the additive genetic effects, permanent environmental effects, and residual effects. The single-step GBLUP (ssGBLUP) approach was implemented, integrating pedigree-based (A-matrix) and genomic-based (G-matrix) relationships into a combined H-matrix (Aguilar et al., 2010) using default settings in BLUPF90+ software (version 2.48; Misztal et al., 2014). Variance components were estimated using average information REML implemented in the BLUPF90+ software. The GEBV were computed using the precondition conjugate gradient algorithm implemented in BLUPF90+.

Genome-wide association analysis was performed using the BLUPF90 family of programs with the ssGWAS approach. The SNP effects of LPNEB and LEDS were estimated using

POSTGSF90 software (version 1.76; Aguilar et al., 2014). The formula used for estimating SNP effects was as stated in Wang et al. (2012):

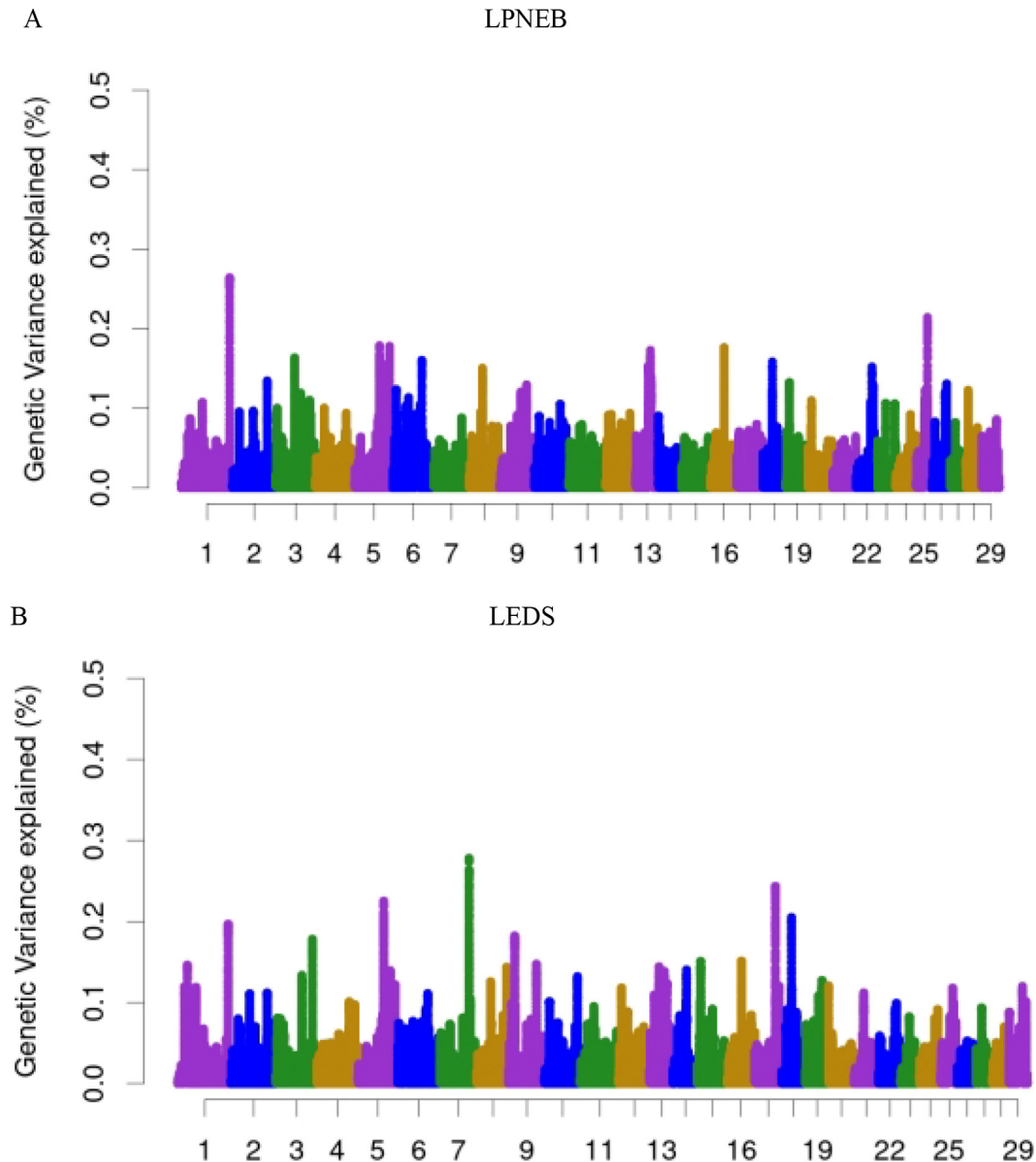
$$\hat{\mathbf{u}} = \mathbf{D}\mathbf{Z}_g' \left[ \mathbf{Z}_g \mathbf{D}\mathbf{Z}_g' \right]^{-1} \hat{\mathbf{a}},$$

where  $\hat{\mathbf{u}}$  is the SNP effect;  $\mathbf{D}$  is the weight matrix of SNPs ( $\mathbf{D} = \mathbf{I}$ ), which means the weight for all SNPs is 1;  $\mathbf{Z}_g$  is an incidence matrix of genotypes for each SNP; and  $\hat{\mathbf{a}}$  is a vector of GEBV for each trait of genotyped animals. The variance of  $i$ th SNP is as follows:  $\text{var}(\hat{u}_i) = u_i^2 \cdot 2p_i(1 - p_i)$ , where  $u_i^2$  is the square of  $i$ th SNP effect and  $p_i$  is the frequency of allele B at SNP  $i$ . A window-based approach was used to account for the combined effects of adjacent SNPs and local linkage disequilibrium structure, thereby improving the estimation of regional genetic variance. The results were presented by the proportion of additive genetic variance explained by each window of 50 adjacent SNPs with an average size of ~240 kb. We used 1 SNP as the moving step of the window, which ensured that we would not miss genomic regions potentially associated with the trait due to the combination of SNPs. The formula for the total additive genetic variance of each window expressed in percentage was as follows:

$$\frac{\text{var}(\hat{\mathbf{a}}_i)}{\sigma_a^2} \times 100 = \frac{\text{var}(\mathbf{Z}_i \hat{\mathbf{u}}_i)}{\sigma_a^2} \times 100,$$

where for all individuals,  $\hat{\mathbf{a}}_i$  is the vector of their breeding values,  $\mathbf{Z}_i$  is the matrix of their SNP content for the SNPs of interest,  $\hat{\mathbf{u}}_i$  the vector of SNP all defined in the  $i$ th genomic region (each window combines 50 consecutive SNPs), and  $\sigma_a^2$  is the total additive genetic variance.

Positional candidate genes and QTL annotations in the top 10 genomic regions were identified using the GALLO R package (Fonseca et al., 2020). The position (coordinate) of selected genomic regions on reference genome assembly UMD3.1 (the used chip version) was converted to the new position (coordinate) on the new reference genome assembly ARS-UCD1.2 through the Lift Genome Annotations tool (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>). Gene Ontology (GO) analyses of the positional candidate genes for each trait were conducted using the g:Profiler website (Kolberg et al., 2023). The selected genomic regions identified for each trait were annotated with Cattle QTLdb (<https://www.animalgenome.org/cgi-bin/QTLdb/BT/index>; Hu et al., 2019). At present, Cattle QTLdb has 192,336 QTL, which were divided into 6 classes, including exterior, production, health, reproduction, milk, meat, and carcass ([https://www.animalgenome.org/cgi-bin/QTLdb/BT/ontrait?class\\_ID=1](https://www.animalgenome.org/cgi-bin/QTLdb/BT/ontrait?class_ID=1)). To avoid the deviation caused by the annotation richness of the different traits, the hypergeometric test approach was adopted for enrichment analysis (Fonseca et al., 2020). In all enrichment analyses (GO and QTL), the Benjamini–Hochberg method was used to correct multiple testing (false discovery rate <0.05). It should be noted that the Cattle QTL dataset currently does not include NEB. All statistical analyses were performed using R (version 4.2.1; R Foundation for Statistical Computing).



**Figure 1.** Additive genetic variance explained by windows of 50 adjacent SNPs across chromosomes for logit-transformed predicted negative energy balance (LPNEB, A) and logit-transformed energy deficiency score (LEDS, B) in early-lactation Holstein cows.

Figure 1 illustrates the distribution of genetic variance explained (%) across the genome for LPNEB and LEDS. Both traits conform to a polygenic inheritance model, with genetic variance contributed by numerous loci distributed across multiple chromosomes. Neither trait is controlled by a single major locus, as the top regions each explain less than 0.5% of the total additive genetic variance, indicating a polygenic genetic architecture. Table 1 compares the top 10 genomic regions explaining the largest proportion of genetic variance for LPNEB and LEDS, highlighting both shared and distinct genetic loci. Several regions, such as those on BTA 1, 5, and 16, contribute to both traits, suggesting a common genetic basis. *Bos taurus* autosomes 5 and 16 were identified in genomic regions with important effects on energy balance (Tetens et al., 2013; Krat-

tenmacher et al., 2019). Additionally, unique genomic regions are associated with each trait, with LPNEB showing specific regions on BTA 6, 13, and 25, and LEDS exhibiting distinct regions on BTA 7, 9, and 17. These differences indicate potential variations in the genetic architecture of the 2 traits. However, we recently reported the overall genetic correlation of 0.85 between LPNEB and LEDS (Hu et al., 2025), suggesting that the common genetic basis between the 2 traits may be dispersed across other genomic regions. Additionally, the 10 biggest genomic regions selected for LPNEB and LEDS only explained a total genetic variance of 1.82% and 1.92%, respectively.

The genes annotation of the selected top 10 genomic regions for LPNEB and LEDS identified a total of 17 positional candidate

**Table 1.** Annotated genes within the top 10 genomic regions explaining the highest proportion of total genetic variance for logit-transformed predicted negative energy balance (LPNEB) and logit-transformed energy deficiency score (LEDS)<sup>1</sup>

Trait	BTA	Position(bp) <sup>2</sup>	Var <sup>3</sup>	Gene <sup>4</sup>
LPNEB	<b>1</b>	<b>150,974,434–1,510,935,985</b>	<b>0.26</b>	<b>PIGP, TTC3</b>
	3	55,153,255–55,383,999	0.16	KYAT3, GTF2B, PKN2
	<b>5</b>	<b>82,540,690–82,638,650</b>	<b>0.18</b>	<b>PPFIBP1</b>
	5	101,175,264–101,308,905	0.16	
	5	111,857,928–112,120,252	0.18	FAM83F, TNRC6B
	6	86,686,814–87,001,138	0.16	ENSBTAG00000015047, UGT2A1, SULT1B1, ENSBTAG00000038214
	13	53,637,893–53,759,622	0.17	SIRPA
	<b>16</b>	<b>44,046,234–44,165,527</b>	<b>0.18</b>	<b>PGD, ENSBTAG00000048790, ENSBTAG00000048747, ENSBTAG00000054239</b>
	18	25,753,024–25,847,046	0.16	
	25	30,079,410–30,209,679	0.21	
LEDS	<b>1</b>	<b>151,023,504–151,147,932</b>	<b>0.20</b>	<b>TTC3,</b>
	3	109,338,317–109,464,393	0.18	GRIK3
	<b>5</b>	<b>82,540,690–82,638,650</b>	<b>0.23</b>	<b>PPFIBP1</b>
	7	88,653,706–88,732,303	0.28	
	9	18,311,181–18,423,364	0.18	
	9	84,820,057–84,931,668	0.15	
	15	10,352,544–10,716,253	0.15	
	<b>16</b>	<b>44,037,583–44,158,583</b>	<b>0.15</b>	<b>PGD, ENSBTAG00000048790, ENSBTAG00000048747, ENSBTAG00000054239</b>
	17	58,683,435–58,772,056	0.24	
	18	25,580,320–25,699,738	0.21	CCDC102A, ADGRG5, ADGRG1

<sup>1</sup>Overlapping genomic regions between LPNEB and LEDS are shown in bold.<sup>2</sup>Position(bp) = Starting and ending coordinates of the genomic region in the UMD3.1 assembly.<sup>3</sup>Var = percentage of genetic variance explained by the genomic regions within the genomic region.<sup>4</sup>The EMBL symbols of annotated genes using the *Bos taurus* ARS-UCD 1.2 assembly ([http://ftp.ensembl.org/pub/release-110/gtf/bos\\_taurus/](http://ftp.ensembl.org/pub/release-110/gtf/bos_taurus/)).

genes for LPNEB and 10 positional candidate genes for LEDS, with 6 positional candidate genes shared between the 2 traits. The GO enrichment analysis for LPNEB and LEDS indicated a significant overlap in metabolic processes (Figure 2A and 2B), particularly in D-gluconate metabolic (GO: 0019521), D-gluconate catabolic (GO: 0046177) and pentose biosynthesis (GO: 0019322), all influenced by the *PGD* gene, suggesting common regulatory mechanisms for energy balance. However, LPNEB is more strongly associated with transferase activity (GO: 0016740; *TTC3*, *KYAT3*, *GTF2B*, *PKN2*, *UGT2A1*, *SULT1B1*, *ENSBTAG00000038214*), indicating a broader metabolic adaptation role. In contrast, LEDS exhibits additional enrichment in neuronal signaling pathways (GO: 0097451, *ADGRG1*; GO: 0097449, *ADGRG1*), suggesting a potential link between metabolic and neurological regulation. These results suggest a shared genetic basis for energy balance, with LPNEB emphasizing metabolic detoxification and enzymatic regulation, and LEDS integrating neuronal signaling into energy homeostasis.

The QTL enrichment analysis of the top 10 genomic regions associated with LPNEB and LEDS against the Cattle QTLdb identified significant associations with multiple economically important traits (Figure 2C and 2D). The QTL enrichment analyses of LPNEB and LEDS revealed 2 shared enriched regions: SCS and interval to first estrus after calving, both on BTA16. These suggest energy metabolic potential has common regulatory mechanisms related to health and fertility (Cole et al., 2011). Beyond these overlaps, LPNEB showed strong enrichment on BTA6 for milk production traits, including milk yield, protein percentage, and kappa-casein content. Previous studies found milk production-related QTL within this genomic region (Buitenhuis et al., 2016; van den Berg et al., 2020; Pedrosa et al., 2021), suggesting a potential genetic link between energy balance and milk composition. In

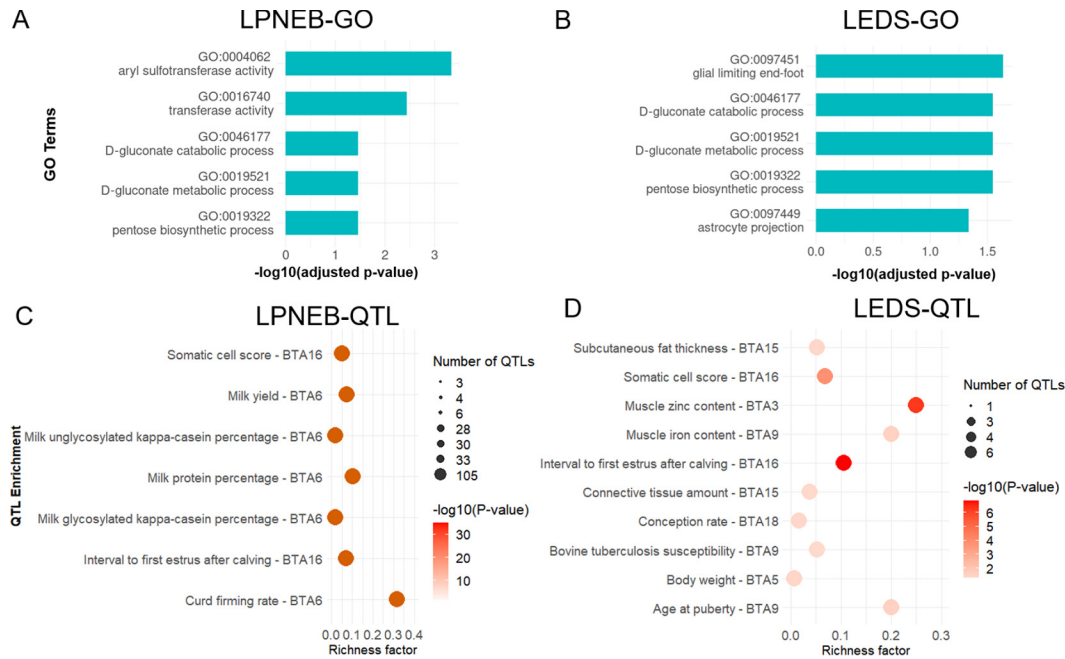
contrast, LEDS was enriched for traits related to muscle development, fat metabolism, immunity, and growth, distributed across multiple chromosomes, such as BTA3, BTA5, BTA9, BTA15, and BTA18. The association with muscle zinc content is particularly noteworthy given its role in protein synthesis, immune function, and overall metabolic efficiency (Hawken et al., 2012; Mateescu et al., 2017). These findings indicate that LPNEB and LEDS share QTL but are also influenced by distinct genetic architectures.

In conclusion, we found that both LPNEB and LEDS exhibited polygenic architectures, but they did not have the same magnitude of genetic variance explained. The presence of shared genomic regions was found to partially explain the high genetic correlation observed between LPNEB and LEDS. However, a more detailed interpretation requires the assessment of local genetic correlations across the genome. Annotation of candidate genes and their functional analysis showed that LPNEB and LEDS were indeed related to energy metabolism in dairy cows. The findings of these QTL analyses indicated that both LPNEB and LEDS are genetically associated with fertility, and immune response (SCS), reinforcing their relevance in cattle breeding for improving production and health outcomes. These findings improve our understanding of the genetic background of LPNEB and LEDS, thereby providing new insights into the mechanisms underlying energy balance in dairy cattle.

## References

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. <https://doi.org/10.3168/jds.2009-2730>.
- Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step





**Figure 2.** The Gene Ontology (GO) terms (A and B) and QTL annotation (C and D) based on candidate genes of logit-transformed predicted negative energy balance (LPNEB) and logit-transformed energy deficit score (LEDS).

- genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. *Proc. 10th World Congr. Genet. Appl. Livest. Prod.* Accessed Aug. 21, 2022. <http://www.ainfo.inia.uy/digital/bitstream/item/15445/1/Aguilar-et-al.-2014.-WCGALP.pdf>.
- Buitenhuis, B., N. A. Poulsen, G. Gebreyesus, and L. B. Larsen. 2016. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post translational modifications in Danish Holstein and Danish Jersey cattle. *BMC Genet.* 17:114. <https://doi.org/10.1186/s12863-016-0421-2>.
- 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>.
- Cole, J. B., G. R. Wiggans, L. Ma, T. S. Sonstegard, T. J. Lawlor, B. A. Crooker, C. Van Tassell, J. Yang, S. Wang, L. Matukumalli, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics* 12:408. <https://doi.org/10.1186/1471-2164-12-408>.
- 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>.
- Fonseca, P. A. S., A. Suárez-Vega, G. Marras, and Á. Cánovas. 2020. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *Gigascience* 9:gaa149. <https://doi.org/10.1093/gigascience/gaa149>.
- 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*. Accessed. <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>.
- 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., A. Vanlierde, F. Dehareng, E. Froidmont, and E. Gplus. Consortium. 2017. Prediction of energy status of dairy cows using MIR milk spectra. *Book of Abstracts of the 68th Annual Meeting of the European Federation of Animal Science*, Tallin, Estonia. Wageningen Academic Publishers.
- 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>.
- Hawken, R. J., Y. D. Zhang, M. R. S. Fortes, E. Collis, W. C. Barris, N. J. Corbet, P. J. Williams, G. Fordyce, R. G. Holroyd, J. R. W. Walkley, W. Barendse, D. J. Johnston, K. C. Prayaga, B. Tier, A. Reverter, and S. A. Lehnert. 2012. Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J. Anim. Sci.* 90:1398–1410. <https://doi.org/10.2527/jas.2011-4410>.
- 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>.
- Hu, Z. L., C. A. Park, and J. M. Reecy. 2019. Building a livestock genetic and genomic information knowledge base through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* 47(D1):D701–D710. <https://doi.org/10.1093/nar/gky1084>.
- Kolberg, L., U. Raudvere, I. Kuzmin, P. Adler, J. Vilo, and H. Peterson. 2023. g:Profiler—Interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* 51(W1):W207–W212. <https://doi.org/10.1093/nar/gkad347>.
- 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>.
- Mateescu, R. G., D. J. Garrick, and J. M. Reecy. 2017. Network analysis reveals putative genes affecting meat quality in Angus cattle. *Front. Genet.* 8:171. <https://doi.org/10.3389/fgene.2017.00171>.
- 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 economi-

- cally important traits in dairy cattle. *J. Dairy Sci.* 100:10251–10271. <https://doi.org/10.3168/jds.2017-12968>.
- Misztal, I., S. Tsuruta, D. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs. Accessed Mar. 3, 2025. [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all8.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all8.pdf).
- Pedrosa, V. B., F. S. Schenkel, S. Y. Chen, H. R. Oliveira, T. M. Casey, M. G. Melka, and L. F. Brito. 2021. Genomewide association analyses of lactation persistency and milk production traits in Holstein cattle based on imputed whole-genome sequence data. *Genes (Basel)* 12:1830. <https://doi.org/10.3390/genes12111830>.
- Pérez-Báez, J., C. A. Risco, R. C. Chebel, G. C. Gomes, L. F. Greco, S. Tao, I. M. Thompson, B. C. do Amaral, M. G. Zenobi, N. Martinez, C. R. Staples, G. E. Dahl, J. A. Hernández, J. E. P. Santos, and K. N. Galvão. 2019. Association of dry matter intake and energy balance prepartum and postpartum with health disorders postpartum: Part I. Calving disorders and metritis. *J. Dairy Sci.* 102:9138–9150. <https://doi.org/10.3168/jds.2018-15878>.
- 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>.
- 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>.
- van den Berg, I., R. Xiang, J. Jenko, H. Pausch, M. Boussaha, C. Schrooten, T. Tribout, A. Gjuvland, D. Boichard, Ø. Nordbø, M. Sanchez, and M. Goddard. 2020. Meta-analysis for milk fat and protein percentage using imputed sequence variant genotypes in 94,321 cattle from eight cattle breeds. *Genet. Sel. Evol.* 52:37. <https://doi.org/10.1186/s12711-020-00556-4>.
- Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W. M. Muir. 2012. Genomewide association mapping including phenotypes from relatives without genotypes. *Genet. Res. (Camb.)* 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>.
- Zachut, M., M. Šperanda, A. M. de Almeida, G. Gabai, A. Mobasher, and L. E. Hernández-Castellano. 2020. Biomarkers of fitness and welfare in dairy cattle: Healthy productivity. *J. Dairy Res.* 87:4–13. <https://doi.org/10.1017/S0022029920000084>.
- S. Franceschini, <https://orcid.org/0000-0001-6298-5149>
- Pauline Lemal, <https://orcid.org/0000-0003-4937-0064>
- Clément Grelet, <https://orcid.org/0000-0003-3313-485X>
- Yansen Chen, <https://orcid.org/0000-0002-8593-4384>
- Katrien Wijnrocx, <https://orcid.org/0000-0001-5518-4933>
- Hélène Soyeurt, <https://orcid.org/0000-0001-9883-9047>
- Nicolas Gengler <https://orcid.org/0000-0002-5981-5509>
- The China Scholarship Council (Beijing, China) 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 “DEEP-SELECT” (grant number T.0095.19).
- This article does not contain any studies with human or animal subjects and did not require IACUC/IRB approval.
- The authors have not stated any conflicts of interest.
- Nonstandard abbreviations used:** EDS = energy deficiency score; GO = Gene Ontology; LEDS = logit-transformed energy deficiency score; LPNEB = logit-transformed predicted negative energy balance; MIR = mid-infrared; NEB = negative energy balance; ssGWAS = single-step GWAS.

## Notes

Hongqing Hu, <https://orcid.org/0000-0002-6408-1845>

Hadi Atashi, <https://orcid.org/0000-0002-6853-6608>