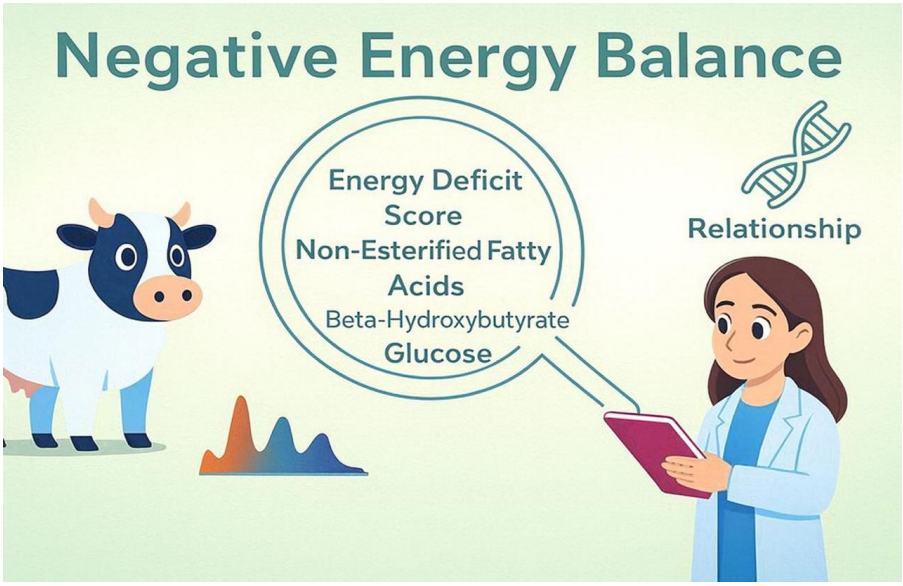
LIÈGE université
Gembloux
Agro-Bio Tech

Investigating the relationships between predicted negative energy balance and its biomarkers in early lactation dairy cows using genetic analysis

Hongqing Hu



Supervisor: Prof. Nicolas Gengler

Year: 2025

COMMUNAUTÉ FRANÇAISE DE BELGIQUE

UNIVERSITÉ DE LIÈGE – GEMBLOUX AGRO-BIO TECH

Investigating the relationships between predicted negative energy balance and its biomarkers in early lactation dairy cows using genetic analysis

Hongqing Hu

Dissertation originale présentée en vue de l'obtention du grade de docteur en sciences agronomiques et ingénierie biologique

Promoteur: Prof. Nicolas Gengler

Année civile: 2025

Abstract

Hongqing Hu. (2025). "Investigating the relationships between predicted negative energy balance and its biomarkers in early lactation dairy cows using genetic analysis". PhD Dissertation in English.

Gembloux, Belgium, Gembloux Agro-Bio Tech, University of Liege.

160 pages, 22 figures, 13 tables.

Summary:

Livestock production constitutes a critical component of global food security, yet it faces persistent challenges of production efficiency, environmental sustainability, and animal health. In high-yielding dairy cows, early lactation is marked by negative energy balance (NEB), when the energy demands of milk synthesis exceed energy intake from feed. This metabolic state induces excessive mobilization of body reserves, disruption of metabolic homeostasis, impaired fertility, and heightened disease susceptibility. These consequences not only impose substantial economic costs but also raise animal welfare concerns, thereby undermining the sustainability and social acceptability of modern dairy systems. Improving energy balance is therefore essential for production efficiency, animal well-being, and long-term industry viability. However, genetic selection against NEB has been constrained by the difficulty of large-scale phenotyping. This thesis aimed to develop and evaluate genetic strategies for characterizing and improving energy balance in dairy cows by integrating large-scale genomic data with innovative proxy phenotypes.

First, measured reference NEB (RNEB) was compared with predicted NEB (PNEB), a novel energy deficiency score (EDS), 15 biomarkers, and 3 production traits, showing moderate to high phenotypic correlations. Genetic parameters of 20 traits were estimated using a 20-traits repeatability model, with heritabilities ranging from 0.16 to 0.38. Genetic correlations between logit-transformed PNEB (LPNEB) and the studied traits varied from –0.60 to 0.87, with the strongest associations for logit-transformed EDS (LEDS) (0.85) and non-esterified fatty acids (NEFA) (0.87). Collectively, the 19 traits explained 89% of the genetic variance of LPNEB, with biomarkers alone accounting for 82%, LEDS alone 65%, and NEFA alone 62%. Recursive modeling further identified eight traits, including NEFA and LEDS, as highly dependent on LPNEB, underscoring their potential as robust biomarkers.

Second, SNP-based genomic correlation analyses, extended to the chromosomal level, revealed strong associations between LPNEB, LEDS, and metabolic biomarkers, particularly NEFA. Independent contribution analyses confirmed NEFA as the

primary driver of both LPNEB and LEDS, while chromosomal scans identified BTA19 and BTA25 as candidate regions underlying NEB regulation.

Third, single-step genome-wide association analyses with a 50-SNP sliding window, using 30,634 records from 25,287 first-parity cows and 566,170 SNPs from 3,757 genotyped animals, identified the top genomic regions for LPNEB and LEDS. Three regions (BTA1, BTA5, and BTA16) were shared between the two traits, alongside unique loci supporting distinct genetic architectures. Candidate gene analyses revealed 17 genes for LPNEB and 10 for LEDS, with 6 shared. Functional enrichment highlighted roles in energy metabolism for LPNEB, and additional neuronal signaling pathways for LEDS. QTL enrichment indicated associations with fertility and somatic cell score.

In conclusion, this thesis explored the genetic correlations between LPNEB, LEDS and multiple proxy traits, metabolic biomarkers, and production traits. The results revealed strong associations with key indicators such as NEFA and identified genomic regions and candidate genes underlying NEB regulation. These findings enhance understanding of the genetic mechanisms of LPNEB and LEDS. Moreover, the results provide evidence supporting the implementation of the MIR-based indicator LEDS as a reliable and scalable tool for improving resilience and energy efficiency in dairy cows through genomic selection and metabolic monitoring.

Résumé

Hongqing Hu. (2025). "Étude génétique des relations entre l'équilibre énergétique négatif prédit et ses biomarqueurs chez les vaches laitières".Thèse de doctorat en anglais.

Gembloux, Belgique, Gembloux Agro-Bio Tech, Université de Liège.
160 pages, 22 figures, 13 tableaux.

Résumé:

La production animale constitue un pilier essentiel de la sécurité alimentaire mondiale, mais elle fait face à des défis persistants liés à l'efficacité de production, à la durabilité environnementale et à la santé animale. Chez les vaches laitières à haut potentiel génétique, le début de lactation est caractérisé par une déséquilibre énergétique négatif (NEB), lorsque les besoins énergétiques liés à la synthèse du lait excèdent l'apport énergétique provenant de l'alimentation. Cet état métabolique entraîne une mobilisation excessive des réserves corporelles, une perturbation de l'homéostasie métabolique, une baisse de la fertilité et une susceptibilité accrue aux maladies. Ces conséquences engendrent non seulement des pertes économiques considérables, mais soulèvent également des préoccupations en matière de bien-être animal, compromettant ainsi la durabilité et l'acceptabilité sociale des systèmes laitiers modernes. L'amélioration de l'équilibre énergétique représente donc un enjeu majeur pour l'efficacité de production, le bien-être animal et la viabilité à long terme du secteur laitier. Toutefois, la sélection génétique contre la NEB reste limitée par la difficulté de phénotypage à grande échelle.

Cette thèse visait à développer et à évaluer des stratégies génétiques pour caractériser et améliorer l'équilibre énergétique chez les vaches laitières, en intégrant des données génomiques de grande ampleur à des phénotypes indicateurs innovants. Dans un premier temps, l'équilibre énergétique de référence mesuré (RNEB) a été comparé à l'équilibre énergétique prédit par spectroscopie moyen infrarouge (MIR) (PNEB), à un nouvel indicateur composite de déficit énergétique (EDS), à quinze biomarqueurs métaboliques et à trois caractères de production, révélant des corrélations phénotypiques modérées à élevées. Les paramètres génétiques de vingt caractères ont été estimés à l'aide d'un modèle de répétabilité à vingt caractères, avec des héritabilités comprises entre 0,16 et 0,38. Les corrélations génétiques entre le PNEB transformé en logit (LPNEB) et les autres caractères variaient de –0,60 à 0,87, avec les associations les plus fortes pour le score d'insuffisance énergétique transformé en logit (LEDS) (0,85) et pour les acides gras non estérifiés (NEFA) (0,87).

Dans l'ensemble, les 19 caractères étudiés expliquaient 89 % de la variance génétique du LPNEB, dont 82 % par les biomarqueurs seuls, 65 % par LEDS seul et 62 % par NEFA seul. Une modélisation récursive a ensuite mis en évidence huit caractères, dont NEFA et LEDS, comme étant hautement dépendants du LPNEB, soulignant leur potentiel en tant que biomarqueurs robustes.

Dans un second temps, les analyses de corrélations génomiques basées sur les SNP, étendues au niveau chromosomique, ont révélé de fortes associations entre LPNEB, LEDS et plusieurs biomarqueurs métaboliques, en particulier les NEFA. Les analyses de contribution indépendante ont confirmé que les NEFA constituaien le principal déterminant des variations observées pour LPNEB et LEDS, tandis que les analyses par chromosome ont identifié les régions candidates BTA19 et BTA25 comme impliquées dans la régulation de la NEB.

Enfin, une étude d'association génomique à une étape (single-step GWAS) réalisée avec des fenêtres glissantes de 50 SNP, à partir de 30 634 enregistrements provenant de 25 287 vaches primipares et de 566 170 SNP issus de 3 757 animaux génotypés, a permis d'identifier les principales régions génomiques associées à LPNEB et LEDS. Trois régions (BTA1, BTA5 et BTA16) étaient communes aux deux caractères, tandis que d'autres loci spécifiques mettaient en évidence des architectures génétiques distinctes. L'annotation fonctionnelle des gènes candidats a révélé 17 gènes pour LPNEB et 10 pour LEDS, dont 6 étaient partagés. L'enrichissement fonctionnel a mis en évidence des voies métaboliques énergétiques pour LPNEB et des voies additionnelles de signalisation neuronale pour LEDS. De plus, l'enrichissement en QTL a indiqué des associations avec la fertilité et le score cellulaire somatique.

En conclusion, cette thèse a exploré les corrélations génétiques entre LPNEB, LEDS et un ensemble de caractères indicateurs, de biomarqueurs métaboliques et de caractères de production. Les résultats ont révélé de fortes associations avec des indicateurs clés tels que les NEFA et ont permis d'identifier plusieurs régions génomiques et gènes candidats impliqués dans la régulation de la NEB. Ces résultats renforcent la compréhension des mécanismes génétiques sous-jacents à LPNEB et LEDS, et soutiennent l'utilisation de l'indicateur LEDS dérivé du MIR comme un outil fiable et évolutif pour améliorer la résilience et l'efficacité énergétique des vaches laitières grâce à la sélection génomique et au suivi métabolique.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to the China Scholarship Council (CSC) for providing the fellowship that made this doctoral research possible. I am also deeply grateful to the University of Liège in Belgium for offering me the opportunity to pursue my doctoral studies and for fostering a supportive, inspiring, and nurturing academic environment throughout this period. This environment has not only greatly enriched my research training, but has also significantly broadened my academic perspectives and contributed to my personal growth.

My deepest gratitude goes to my supervisor, Prof. Nicolas Gengler, for his exceptional mentorship, scientific rigor, and constant encouragement throughout this journey. His insightful guidance and remarkable patience have not only shaped the quality of this thesis, but have also profoundly influenced my way of thinking as a researcher. I am especially grateful for the way he taught me to work independently, think critically, and take responsibility for my own scientific decisions. I sincerely appreciate the countless discussions, constructive feedback, and invaluable opportunities he has provided each of which has played a central role in my academic development.

I would also like to express my sincere appreciation to the members of my thesis committee Prof. Yves Beckers, Dr. Frédéric Dehareng, Prof. Nicolas Gengler, Prof. Bingjie Li, Prof. Martine Schroyen, Prof. Hélène Soyeurt, and Dr. Katrien Wijnrocx for their valuable comments, insightful suggestions, and generous availability throughout the evaluation of my work. I am deeply honored by the time and expertise each of you has contributed to my doctoral journey. My heartfelt thanks also go to the jury members for their kind participation in evaluating my thesis and for the thoughtful questions that helped refine and strengthen my research perspective.

I would also like to express my heartfelt gratitude to Dr. Grelet and Sébastien Franceschini. I am especially grateful to Dr. Grelet for his continuous support throughout my research, particularly for his expertise and guidance in the acquisition, processing, and interpretation of MIR spectral data, as well as for the valuable advice he provided during each annual PhD committee meeting. I am also thankful to Sébastien Franceschini for providing the data and offering technical assistance, which laid an important foundation for the progress of this study. Their expertise and support have been truly invaluable.

I would like to express my sincere gratitude to my colleague Dr. Hadi for his continuous support, valuable advice, and generous help throughout my research. His

expertise, detailed feedback, and patient guidance have greatly improved the quality of my work. His encouragement and positive attitude have supported me during difficult moments, and his help has contributed significantly to my academic growth, research thinking, and personal development.

I would also like to express my sincere thanks to my colleagues in the laboratory. Alice Markey, Bérat Hugo, Hélène Wilmot, Yansen Chen, Pauline Lemal, Katrien Wijnrocx, Sylvie Vanderick, Darlene Ana Souza Duarte, and Marie Olivier. Your collaboration, discussions, and support have played an important role throughout my research. Whether through scientific exchanges, sharing technical expertise, or helping with experiments and data processing, each of you has contributed greatly to the progress of my work and has made daily research more efficient and enjoyable. Thank you for your professionalism, patience, and kindness, which have created a productive and pleasant working environment. I feel very fortunate to have worked with all of you.

I am sincerely grateful to the Chinese Students and Scholars Association (CSSA) in Gembloux for making my journey abroad less lonely and for bringing warmth and support throughout my time in Belgium. I would also like to thank my friends at Gembloux Agro-Bio Tech. Axelle and her family, Kambiré Sami Joseph, Esther, Lily, Bannour Yosra, Yuling Ma, Minmin Xie, Mei Ren, Yelin Zheng, Yumei Hu, Xin Jiang, Xuemei Zhao, Lei Zhang, Huiyuan Li, Shu Wang, Yueling Yang, Ning Ding, Shuang Wang, Guangchun Song, Pengxiang Yuan, Wei He, Yining Xie, and others for their kindness, encouragement, and companionship. Your presence made my life in Gembloux warmer, happier, and much less difficult. The support, laughter, and moments we shared together have been a great comfort, and being able to walk this journey with you has been one of the greatest blessings of my years abroad.

During my time in Gembloux, the birth of my child, Jiarui, became one of the most beautiful and unforgettable moments of my life. His arrival brought light and warmth to my days abroad and gave me strength during many challenging moments. I am deeply grateful to my mother, who came all the way from home to support me with love and care, allowing me to focus on both my studies and my new role as a mother. I am also sincerely thankful to my husband, Yansen, whose steady support, understanding, and quiet dedication helped me face the demands of studying abroad and caring for a newborn with confidence and peace. I am grateful as well for the kindness of my landlady and for Mengyu's companionship, which made this special period of my life much warmer.

My heartfelt gratitude also goes to my family. To my parents, thank you for your unconditional love and constant support; your encouragement has sustained me

through every stage of this journey. To my younger brother, thank you for your understanding and quiet support. I am also deeply grateful for the love and guidance of my grandparents, whose teachings from my childhood have shaped the person I am today. My family has always been my source of strength, comfort, and motivation. Your love and support have been fundamental to my ability to complete this thesis, and I am profoundly thankful.

12/2025, Gembloux

致谢

大鹏一日同风起，扶摇直上九万里。

余自东方来，乘长风、渡重洋，跨山海千重，越云水万里。初至比国，让布卢小城烟雨如画，青石为阶、古木生香，虽异乡，亦若桃源。博士数载，风雨兼程；清晨踏露而往，夜半对灯而归。或与数据鏖战，似冲霄汉；或与难题交锋，若斗群峰。

寒夜漫长，唯有孤灯相伴；

道路崎岖，却见星辰作引。

而今执笔回望，胸中浩然之气如潮涌，天地山河皆来助我；感恩之心如长风，拂过岁月、荡尽尘埃。此生能得此行，纵万般辛劳，亦值得三分豪情、七分热血。故作此文，以谢群贤，以慰亲友，以铭此程，以记此心。

首当感谢中国国家留学基金管理委员会 (CSC)。若无其资助，则我之舟难度沧海；若无其托举，则我之梦难凌云。千山万水之间，是祖国为我架桥铺路，使我得以离乡千里、追问学术之源。此恩如春雷开霁，振我胸臆；亦如大河奔流，托我远行。列日大学张臂相迎，堂宇肃穆，群贤毕至，其学术之风如清泉入心，其包容之气如长风拂面。在此求学，仿佛登高揽岳，目极千里，心境自开；得以与世界学者论道，与学术前沿并行，是此生难得之际遇。此恩如高山巍峨，朗日长明，照我前路，温我此心。

吾师 Nicolas Gengler 教授，实为我学旅中的擎天一柱，其学问如瀚海浩渺，其见识如星辰灿然；与之论学，若清风入怀，拂尽胸臆之疑；若明月照心，万象随之朗然。人生路上每当迷惘，他一言即破千重雾障；学途之中每遇艰难，他一念便托举万里山河。其胸怀宽厚，严而有度；其风骨峻洁，柔中带刚，使我于求索之境中得以日渐登高、眼界渐阔。能得此师指路，乃我学途鸿福，亦是命中不期而遇之光，三生有幸，一言不虚。

诸位论文委员会的师长——Yves Beckers 教授、Frédéric Dehareng 博士后、Nicolas Gengler 教授、Bingjie Li 教授、Martine Schroyen 教授、Hélène Soyeurt 教授，以及 Katrien Wijnrocx 博士后，皆如北斗明灯，高悬学海之巅，照亮我稿中幽微；其评述精深，若雷霆破雾；其指点入微，如春雨润心，使论文更稳、更深、更明。群贤并立，智慧奔涌，如江海潮生；恩泽之厚，如山岳之重，非数语所能尽述。得蒙诸师审阅与教诲，是我学途之大幸，亦是人生之光。

Dr. Grelet，乃 MIR 光谱之善导、数据海洋之舟楫，其学识如皓月临川，其心境若长风入海；其言简而意深，其思清而境阔，每与之论学，皆如开雾登峰、行云入画。凡遇难题，他总能一语破迷津，使繁者归简、晦者成明；每当我在数据之海中漂泊无依，他的指引便若乘银河上溯，见群星璀璨，令胸中万象顿

时明朗。其助益之广，其启迪之深，贯穿我整个学途，如灯照长夜，如舟渡重洋，使我于学海沉浮之间始终不失方向。胸中万象，常承其启。

Hadi 博士后，于余则如长风挟劲草，暖我研途、助我登攀；其心诚，其志厚，迷惘时为我指途，低落处为我撑船，使我于曲折之中仍得见青天。而 Yansen 博士后亦如长夜之炬、冬境之阳，以笃定与温柔托我稳行；其理解若春风散霜，其陪伴如磐石镇川，使我在异乡求学与人生风浪之间皆能从容前步。二人于我旅途，一为同行良朋，一为生命所依；得与君相识，实为幸事，得与君同行，更为命中所赐之福。

实验室诸贤——Atashi Hadi、Alice Markey、Bérat Hugo、Chen Yansen、Lemal Pauline、Hélène Wilmot、Kartin Wijnrocx、Sylvie Vanderick、Sébastien Franceschini、Souza Duarte Darlene Ana，Marie Olivier 皆为我学途中的山水知音；或与我论学，如激浪拍岸，惊雷破空，思想交锋间涌起千层浪；或与我协作，如清风并行，步调相和，心意相契，使繁复之事若行云流水、轻松自若。诸君或以才识启我眼界，或以善意暖我行程，使异国的实验室不仅有数据与仪器，更有人情与光亮。同行之人，皆为山河；因有诸位，路不再孤，行愈见远。

亦谢让布卢中国学生学者联合会（CSSA），漂泊之途因你们而不孤，异乡风雨因你们而有依，所予之暖，长存胸臆，不负流年。至于在让布卢农学生命科学学院（Gembloux Agro-Bio Tech）相伴的良友——Axelle 以及其家人、Kambiré Sami Joseph、Esther、Lily、Bannour Yosra、马玉玲、解敏敏、任玫、郑宏艳、郑润田、陈风、胡玉梅、蒋欣、赵雪梅、张磊、李慧园、王术、杨越玲、李自馨、丁宁、王凯、段文静、王爽、宋光春、袁鹏翔、谢乙宁、何维……诸君皆为我异国岁月里的清风朗月，共度冷暖，同行风雨；或陪我破难关，或与我共深谈，欢笑满径，步步皆暖。与君相聚，情谊如酒，初入口或烈似风霜，却回味绵长甘醇；举杯可对明月，光照千里；低头亦见温柔，一念即暖。此情若星火，虽起于微芒，却能照亮长夜；照人亦照己，使漂泊之身不孤，使荆棘之路亦有光。得诸友相扶同行，实乃我异国学途之大幸。

值我旅居让布卢之年，欣逢嘉瑞诞生，如春光入怀，使异乡岁月倍添明亮。他的到来，如晨光破晓，照亮前路，使平凡日月皆添温柔与色彩。母亲远道奔赴，以柔情为灯、以慈心为盾；夫君焱森沉稳如山、深情如海，为我托举天地，使我纵遇风浪亦能安然自持。房东之善意如冬日暖阳；梦玉之陪伴如清风入怀，使异国他乡亦生故土之暖。此番恩情，皆铭刻于我心，珍而不忘。

家人，是我行路的灯塔。父母厚德如江河，以无私大爱托我远行；父亲的话语如清风振心，使我在踟蹰之时再次点亮心火；母亲以坚韧和温情护我周全，使我行至风雨处，胸中依旧光芒自持。弟弟性情沉静，却以理解与支持同我并肩，使我在风雨之间胸有长风。亦深感念爷爷奶奶，自幼教我读书明理，以淳厚家风立我根基；此恩深植心底，如根深叶茂，伴我一路成长。家人之恩，深

重无量，是我行远路的底气，是我望星河的依凭。心有所爱，路自生光；一切奔赴皆有源，一切坚持皆有因。愿未来岁月皆顺遂；愿所爱之人皆安康；愿赠我力量与温暖的家人，永沐朗月清风，长携喜乐随身。

亦愿轻落一笔，以谢自己：这些年风雨兼程、跌宕起伏，是我以一腔孤勇撑过漫长寒夜，也是我在无数清晨与深夜之间独自拾起希望；困顿时未曾放弃，迷惘时仍在前行，纵有艰难，却终走至今日。愿往后的我，依旧心怀晴光，仍存鸿翼之志，以坚定与善良迎向每一程山海。

腊月 · 二〇二五
于让布卢（Gembloux）
胡鸿晴
晴光作笔 · 鸿翼为舟 · 书此长章

Table of contents

List of figures

List of tables

List of abbreviations

ACE	Acetone
All_BTA	Combined results across all bos taurus autosomes
AI-REML	Average information restricted maximum likelihood
B_BHBA	Blood β-hydroxybutyrate acid
BCS	Body condition score
BHBA	β-hydroxybutyrate acid
BLUP	Best linear unbiased prediction
BTA	Bos taurus autosome
BW	Body weight
C10:0	Decanoic acid
C14:0	Myristic acid
C16:0	Palmitic acid
C18:0	Stearic acid
C18:1 cis-9	Oleic acid
CIT	Citrate
CI	Consumption index
DE	Digestible energy
DGV	Direct genomic value
DIM	Days in milk
DNA	Deoxyribonucleic acid
DMI	Dry matter intake
DMI2	Predicted dry matter intake
EB	Energy balance
EBV	Estimated breeding value
EDS	Energy deficiency score
EM-REML	Expectation–maximization restricted maximum likelihood
EI	Energy intake
FA	Fatty acids
FDR	False discovery rate
FE	Fecal energy
FP	Fat percentage
FPR	Fat-to-Protein Ratio
FY	Fat yield
GasE	Gaseous energy
GE	Gross energy
GEBV	Genomic estimated breeding value
GH	Growth hormone

GLU	Glucose
GO	Gene ontology
GWAS	Genome-wide association study
h^2	Heritability
HI	Heat increment
HSL	Hormone-sensitive lipase
ICAR	International committee for animal recording
IGF-1	Insulin-like growth factor 1
IRB	Institutional review board
LACE	Log10-transformed milk acetone
Lactof	Lactoferrin
LB_BHBA	Log10-transformed blood β-hydroxybutyrate acid
LCFA	Long-chain fatty acids
LEDS	Logit-transformed energy deficiency score
LIGF-1	Log10-transformed insulin-like growth factor 1
LM_BHBA	Log10-transformed milk β-hydroxybutyrate acid
LPNEB	Logit-transformed predicted negative energy balance
LRNEB	Logit-transformed measured reference negative energy balance
M_BHBA	Milk β-hydroxybutyrate acid
MAF	Minor allele frequency
MCFA	Medium-chain fatty acids
ME	Metabolizable energy
MeanAbsCorr	Mean absolute genomic correlation
MIR	Mid-infrared
MTM	Multi-trait model
MY	Milk yield
NE	Net energy
NEB	Negative energy balance
NEFA	Nonesterified fatty acids
NUE	Nitrogen use efficiency
PCA	Principal component analysis
PEB	Predicted energy balance
PNEB	Predicted negative energy balance
PP	Protein percentage
PY	Protein yield
QTL	Quantitative trait locus
r	Correlation value
R^2	Coefficient of determination

R^2cv	Cross-validation coefficient of determination
REB	Reference energy balance
REML	Restricted maximum likelihood
RFI	Residual feed intake
RM	Recursive model
RMSE	Root mean squared error
RNEB	Reference negative energy balance
RNA	Ribonucleic acid
SCFA	Short-chain fatty acids
SCS	Somatic cell score
SD	Standard deviation
SE	Standard error
SEM	Structural equation models
SNP	Single nucleotide polymorphism
ssGBLUP	Single-step genomic best linear unbiased prediction
ssGWAS	Single-step genome-wide association study
TAG	Triacylglycerol
TCA	Tricarboxylic acid
UE	Urinary energy
VIF	Variance inflation factor
VLDL	Very-low-density lipoprotein
σ^2a	Additive genetic variance
σ^2e	Residual variance
σ^2p	Permanent environmental variance
λ	Structural coefficient

1

Chapter I General introduction

1. Foreword

In this chapter, we define negative energy balance (NEB) in dairy cows and summarize its physiological mechanisms and impacts. We review current methods for assessing NEB, with emphasis on mid-infrared (MIR) spectroscopy-based prediction, and offer insights into energy transformation in the cow body. Finally, we outline the research objectives, which are to explore the relationships between logit-transformed predicted NEB (LPNEB), logit-transformed novel energy deficiency score (LEDS) and key metabolic and production biomarkers, and to investigate the genomic architecture of LPNEB, including the identification of key genomic regions and candidate genes.

2. Definitions of energy balance and negative energy balance in dairy cows

2.1. Energy balance

Energy balance (EB) in dairy cows is the difference between the metabolizable energy (ME) obtained from feed and the total energy required for maintenance, milk production, reproduction, activity, and the deposition or mobilization of body tissues (Thorup et al., 2012). Energy intake is primarily derived from the metabolizable fraction of dry matter intake (DMI). The components of energy expenditure in dairy cows. These components are defined according to established nutritional principles (NRC, 2001):

Maintenance requirements: the basal metabolic functions necessary to sustain life.

Lactation requirements: the energy required for milk synthesis and secretion.

Reproductive and activity costs: the energy used for estrous cycles, pregnancy, and locomotion.

Tissue deposition or mobilization: the energy stored in or released from body reserves.

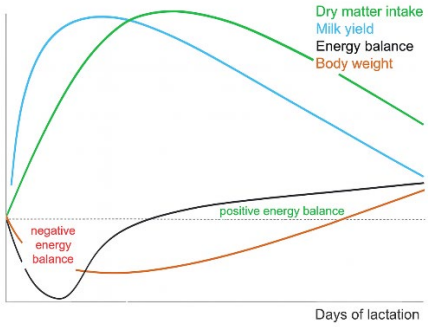


Figure 1-1. Typical lactation curves of milk yield, dry matter intake, energy balance, and body weight in dairy cows (Harder et al., 2019).

2.2. Energy transformation pathway

In dairy nutrition, the flow of energy from feed to the cow's metabolic processes follows a hierarchical transformation (NRC, 2001):

Gross Energy (GE): total chemical energy in feed.

Digestible Energy (DE): energy absorbed after subtracting fecal losses:

$$DE = GE - FE$$

Metabolizable Energy (ME): DE minus urinary and methane losses:

$$ME = DE - (UE + GasE)$$

Net Energy (NE): ME minus heat increment, representing the fraction directly available for maintenance, lactation, growth, and reproduction:

$$NE = ME - HI$$

Overall relationship:

$$NE = GE - FE - UE - GasE - HI$$

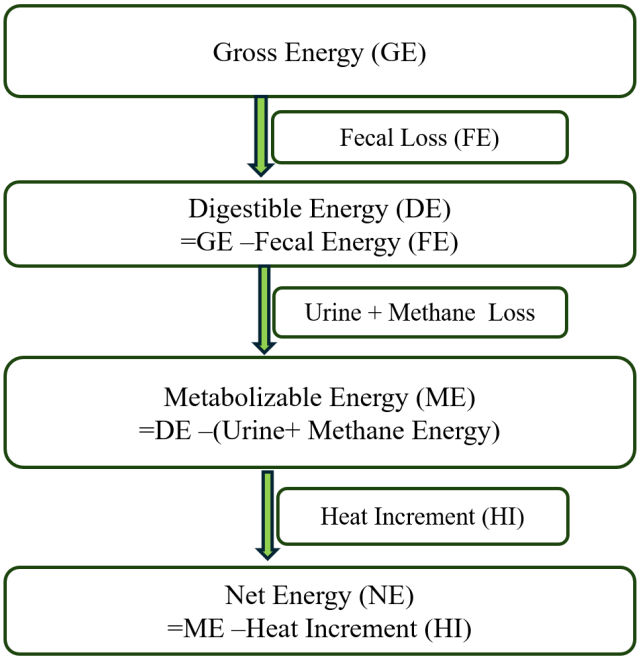


Figure 1-2. Stepwise conversion of feed energy and energy losses at each stage (NRC, 2001).

2.3. Negative energy balance

The EB of dairy cows is defined as the difference between dietary energy intake and total energy expenditure (NRC, 2001):

$$EB = \text{Energy intake} - \text{Energy expenditure}$$

EB = 0: Energy equilibrium dietary energy intake equals the cow's total energy requirements.

EB > 0: Positive energy balance dietary energy intake exceeds requirements; surplus energy supports body tissue accretion and recovery of body condition.

EB < 0: Negative energy balance dietary energy intake is insufficient to meet requirements, leading to mobilization of body fat and protein reserves.

3. Impacts of negative energy balance on health, reproduction, and production

3.1. Health impacts

NEB during early lactation has been linked to increased metabolic stress in dairy cows, contributing to excessive adipose tissue mobilization and higher circulating NEFA concentrations. (Mellouk et al., 2019; Drackley, 1999). When the hepatic capacity for fatty acid oxidation and very low-density lipoprotein (VLDL) export is exceeded, triglycerides accumulate in the liver, increasing the risk of hepatic lipidosis (White, 2015). In parallel, incomplete fatty acid oxidation elevates ketone body production, resulting in higher BHBA concentrations and a greater incidence of clinical and subclinical ketosis (Duffield et al., 2009; McArt et al., 2012).

Beyond metabolic disorders, NEB compromises immune competence. Energy deficiency reduces leukocyte proliferation, impairs neutrophil phagocytic activity, and alters cytokine profiles, thereby weakening both innate and adaptive immune responses (Suriyasathaporn et al., 2000; Ingvartsen and Moyes, 2013). This immunosuppression increases susceptibility to periparturient infections, particularly mastitis and metritis, which in turn exacerbate metabolic stress and delay recovery (Esposito et al., 2014).

3.2. Reproductive impacts

NEB in early lactation has a profound influence on reproductive performance in dairy cows. During NEB, reduced glucose and insulin levels, along with elevated NEFA and BHBA, disrupt the hypothalamic–pituitary–ovarian axis, leading to delayed resumption of ovarian cyclicity after calving (Butler, 2003; Wathes et al., 2007a). These metabolic changes are associated with prolonged anovulatory periods, reduced expression of estrus, and impaired follicular development (Leroy et al., 2006).

Cows experiencing severe NEB often exhibit lower conception rates, increased days open, and extended calving intervals (Walsh et al., 2011). The mobilization of body fat reserves increases the risk of subclinical and clinical ketosis, which further compromises reproductive function through negative effects on oocyte quality and uterine health (Bisinotto et al., 2012; Mekuriaw, 2023). Overall, NEB induced reproductive inefficiencies represent a major challenge for both herd fertility management and farm profitability.

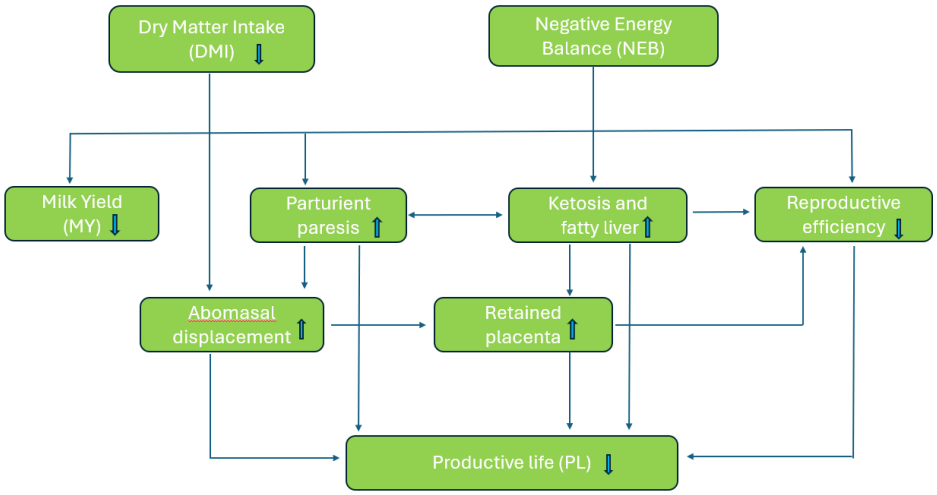


Figure 1-3. Effects of reduced dry matter intake (DMI) and negative energy balance (NEB) on health, reproduction, and productive life in dairy cows (Mekuriaw, 2023).

3.3. Production impacts

NEB during early lactation has profound effects on dairy cow productivity. Cows experiencing NEB often exhibit a decline in body condition score (BCS) due to mobilization of adipose and muscle tissue to meet energy demands (Casaro et al., 2024). Excessive BCS loss (>0.5-1.0 units within the first month postpartum) is associated with reduced peak milk yield and a faster post-peak decline in production (Singh and Bhakat, 2022).

From a lactation curve perspective, NEB may delay the achievement of peak yield and shorten lactation persistency, leading to lower cumulative milk output (Harder et al., 2019). This is partly due to metabolic adaptations during NEB that prioritize survival and homeostasis over maximal milk synthesis. Moreover, severe NEB has been linked to altered milk composition, including reduced protein percentage and changes in milk fatty acid profile, which can affect processing quality and market value (Ducháček et al., 2020). High concentrations of milk long-chain fatty acids and

elevated fat to protein ratio (FPR) are often used as indirect indicators of NEB (Churakov et al., 2021).

At the farm level, these production losses translate into economic consequences, as reduced milk yield and altered composition directly lower revenue, while increased health and reproduction problems raise costs, thereby diminishing overall profitability and sustainability (Sundrum, 2015).

4. Methods for assessing negative energy balance in dairy cows

Accurate evaluation of EB is essential for understanding NEB in dairy cows and for maintaining the metabolic health of high-yielding animals. Although EB can theoretically be defined as the difference between dietary energy intake and energy expenditure, practical applications in research and herd management are primarily focused on its negative state, i.e., NEB. Reliable quantification of NEB not only provides insights into nutrient partitioning and physiological adaptations but also underpins improvements in reproductive performance, disease resistance, and overall herd profitability. Methods for assessing NEB are generally categorized into direct measurements and indirect indicators, which differ substantially in accuracy, cost-effectiveness, feasibility, and suitability for implementation in large-scale dairy production systems.

Table 1-1. Classification and characteristics of methods for assessing negative energy balance in dairy cows.

Classification	Method	Key Indicators	Advantages	Limitations	Reference
Direct Measurement	Calorimetry / Energy Balance Calculation	Energy intake – Energy output (maintenance, lactation, growth, gestation)	Gold standard; provides quantitative estimates of energy deficit	Labor-intensive, requires precise input data; not practical for on-farm application	Huhtanen and Bayat, 2025
Indirect Measurements	Body condition measures	Body condition score (BCS), Body weight	Low-cost, noninvasive, reflects long-term energy status	BCS is subjective; body weight changes lag behind energy status	Bastin et al., 2007 Roche et al., 2009
	Blood metabolites	Nonesterified fatty acids (NEFA), β-hydroxybutyrate (BHB), Glucose	High sensitivity for early detection of NEB; useful for predicting ketosis and fatty liver	Invasive; requires laboratory analysis; limited applicability under field conditions	Zachut et al., 2020; Pires et al., 2022
	Milk-based indicators	Fat-to-protein ratio (FPR)	Easily obtainable during milking; suitable for herd-level screening	Affected by diet, stage of lactation, and mastitis; limited diagnostic accuracy when used alone	Mekuriaw, 2023
	Rumination & activity sensors, automated behavior tracking	Rumination time, activity level, feeding/lying behavior	Noninvasive and real-time monitoring; suitable for large-scale automation	High technological cost; algorithms still require optimization and validation	Elischer et al., 2013

4.1. Direct measurement of negative energy balance

The metabolic chamber technique represents the gold standard for quantifying energy balance in dairy cows (Huhtanen and Bayat, 2025). This method integrates precise measurements of feed intake, fecal and urinary energy losses, and heat production derived from respiratory gas exchange (O_2 consumption, CO_2 and CH_4 output), with heat production commonly calculated using the classical Brouwer equation (Brouwer, 1965). By combining metabolic heat production with milk energy output and changes in body tissue reserves, the technique yields the most accurate

estimation of metabolizable energy use and, consequently, the cow's energy balance. Despite its high accuracy, the metabolic chamber system is extremely expensive to construct and operate, requires intensive labor and specialized technical expertise, and allows only a very limited number of animals to be evaluated at a time. These limitations make the method unsuitable for large-scale phenotyping and therefore impractical for genetic analyses that rely on thousands of animals and repeated measurements. As a result, metabolic chambers are primarily used for metabolic research, calibration studies, and validation of alternative indicators of energy balance.

Direct measurement of NEB in dairy cows is derived from the general concept of EB, which is defined as the difference between dietary energy intake and the energy expended for maintenance, milk production, activity, growth, and reproduction (Banos and Coffey, 2010). Although conceptually accurate, this approach demands precise, individual level data on both feed intake and energy expenditure, most of which, apart from milk yield, are difficult to quantify under practical field conditions (Thorup et al., 2012). In addition, the method is subject to considerable estimation errors, which restricts its feasibility for routine implementation in large-scale commercial dairy herds.

4.2. Indirect measurement of negative energy balance

4.2.1 The role of body condition score in monitoring dairy cow negative energy balance

Among the phenotypic indicators used to assess NEB in dairy cows, BCS is the most widely applied method. The assessment of BCS is a non-invasive, simple, and cost-effective approach that evaluates subcutaneous fat reserves through visual appraisal and tactile evaluation (Roche et al., 2009). Different systems are used across the world. For example, Wildman et al. (1982) proposed a system where dairy cows are given a BCS based on a 5-point scoring system (1–5) with quarter points. Various other BCS systems have been developed across the world, such as a 6-point scale (0–5) first proposed in the United Kingdom (Mulvany, 1977). Especially in the context of genetic evaluations a 9-point scale with unit increments (1–9) is commonly used, also in the Walloon Region of Belgium (Bastin et al., 2007). Regardless of the scale used to measure BCS, low values always reflect emaciation, and high values equate to obesity. It is particularly relevant during the transition and early lactation periods, when dairy cows are at high risk of NEB due to increased energy demands. In these stages, BCS provides an effective reflection of fat mobilization and long-term energy status (Roche et al., 2009). Numerous studies have demonstrated strong associations between BCS and major performance traits such as milk yield (Dezetter et al., 2024), reproductive performance (Roche et al., 2007), and metabolic health (Gillund et al., 2001), making it a valuable proxy for identifying cows at risk of NEB. However, BCS has inherent limitations: it is subjective, subject to inter-observer variability, and

relatively insensitive to short-term fluctuations in energy balance. Despite these limitations, BCS remains a widely used tool in routine herd management and can be effectively combined with other objective indicators to improve the accuracy and timeliness of NEB monitoring.

4.2.2 Biomarkers for monitoring negative energy balance in dairy cows

Biomarkers commonly used to monitor NEB in dairy cows include blood metabolites such as NEFA, BHBA, glucose, insulin, and leptin, as well as the milk FPR (Mekuriaw, 2023). These indicators provide valuable information on fat mobilization, energy supply, and metabolic regulation status, and have been extensively applied in both experimental studies and herd level management (Caixeta and Omontese, 2021; Churakov et al., 2021). Nevertheless, their application is constrained by several limitations: blood sampling is invasive, laboratory assays are costly, and many biomarkers are highly sensitive to management and environmental conditions. Furthermore, most biomarkers characterize NEB retrospectively rather than prospectively, limiting their value as early-warning indicators. Therefore, the development of stable, rapid, cost-effective, and predictive biomarkers remains a critical research priority in modern dairy production systems.

4.2.3 Mid-Infrared spectroscopy as a tool for negative energy balance monitoring in dairy cows

The evaluation of NEB in dairy cows has traditionally relied on directly measured biomarkers, such as NEFA, BHBA, and insulin-like growth factor 1 (IGF-1), which provide reliable indicators of energy mobilization and metabolic stress (Xu et al., 2020; Pires et al., 2022). Although physiologically informative, these biomarkers require invasive sampling procedures (e.g., blood collection) and costly laboratory analyses, which limit their applicability for large-scale or routine herd monitoring.

Mid-infrared (MIR) spectroscopy has long been a successful and well-established technique for analyzing major milk components. As a rapid, non-invasive, and cost-effective technology, MIR has been routinely applied in milk recording for decades to determine milk composition, including fat, protein, and lactose (Gengler et al., 2016). Pioneering work by Soyeurt et al. (2006) further demonstrated that MIR spectra can be calibrated to predict the fatty acid composition of milk, providing an efficient alternative to gas chromatography. Building on this foundation, McParland et al. (2015) showed that MIR spectra can also predict energy intake (EI) and energy balance (EB), with moderate heritability and strong genetic correlations with measured traits, thus confirming their potential as proxies for true EB in genetic improvement programs. These findings collectively highlight the strong potential of MIR-based predictions as a promising tool for estimating EB in dairy cows.

More recently, MIR spectroscopy has been extended to predict a broad range of metabolic biomarkers associated with NEB, including NEFA, BHBA, glucose, and

citrate (Grelet et al., 2021). By exploiting spectral variation in routinely collected milk samples, MIR-based prediction equations can generate proxies for energy related traits in a high-throughput, non-invasive, and cost-effective manner. Importantly, MIR predictions enable repeated longitudinal measurements across thousands of animals, making them particularly suited for genetic analyses and herd-level NEB monitoring.

Despite these advantages, several challenges remain. Prediction accuracy is influenced by lactation stage, diet, and calibration strategy, which may reduce the robustness of MIR-based models across herds and environments. In addition, indirect indicators such as the FPR are prone to confounding effects, limiting their specificity for NEB assessment (Buttchereit et al., 2011). Furthermore, because FPR is a ratio trait, changes in its value cannot distinguish whether fat or protein is responsible. Since both fat and protein are key economic traits, reducing either component to modify FPR is not a feasible or desirable strategy for dairy producers. Consequently, improving calibration procedures, expanding and diversifying reference datasets, and integrating MIR predictions with complementary phenotypic and genomic information are essential steps to enhance the accuracy, robustness, and applicability of MIR-based NEB monitoring tools in modern dairy production systems.

5. Assessment of negative energy balance using biomarkers

5.1. Endocrine and metabolic adaptations to negative energy balance

To cope with the pronounced energy deficit of early lactation, dairy cows undergo a cascade of coordinated endocrine and metabolic adaptations. The insulin-to-glucagon ratio decreases, while secretion of growth hormone (GH) increases, stimulating lipolysis in adipose tissue through activation of hormone-sensitive lipase (HSL) (Dietz and Schwartz. 1991). This results in the release of NEFA and glycerol into circulation. Glycerol serves as a substrate for hepatic gluconeogenesis to sustain blood glucose levels, whereas NEFA is primarily oxidized in the liver, producing acetyl-CoA. When acetyl-CoA accumulates beyond the capacity of the tricarboxylic acid (TCA) cycle, it is redirected toward ketogenesis, leading to the formation of ketone bodies such as BHBA.

These endocrine and metabolic shifts not only ensure short-term energy supply for lactation but also give rise to a range of measurable biomarkers in blood and milk. Elevated NEFA and BHBA, reduced glucose and IGF-1, as well as alterations in milk composition and fatty acid profiles, are direct reflections of the cow's metabolic status. As such, they provide valuable indicators for monitoring NEB at both the individual and herd level.

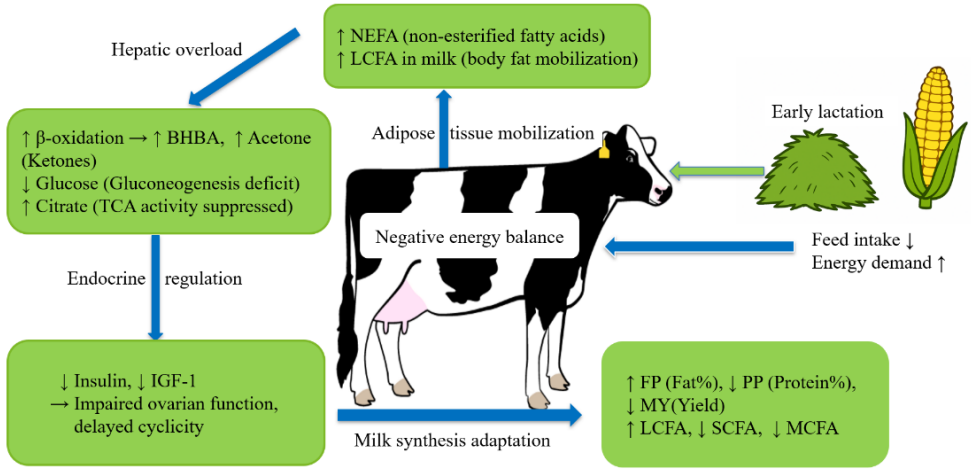


Figure 1-4. Conceptual diagram illustrating the physiological mechanisms underlying negative energy balance in early-lactation dairy cows (Kaniamuthan et al., 2025).

5.2. Blood biomarkers associated with negative energy balance

NEB in early-lactating dairy cows is closely reflected in alterations of blood metabolites and hormones, particularly NEFA, BHBA, acetone, and IGF-1 (Xu et al., 2020; Zachut et al., 2020; Pires et al., 2022). Elevated NEFA indicates enhanced mobilization of adipose reserves, while excessive hepatic β-oxidation of NEFA results in the accumulation of ketone bodies, mainly BHBA and acetone, which are strongly associated with ketosis. In contrast, circulating IGF-1 concentrations decrease during NEB, reflecting impaired hepatic synthesis and altered endocrine regulation, and this reduction has been linked to reduced fertility in dairy cows (Wathes et al., 2007b).

Among these indicators, NEFA is one of the most established blood biomarkers for assessing the metabolic status of dairy cows. Originating predominantly from adipose tissue mobilization, NEFA release is markedly enhanced when feed intake fails to meet the energy demand of lactation. Under NEB, decreased insulin and elevated growth hormone (GH) stimulate hormone-sensitive lipase (HSL), accelerating lipolysis and the release of NEFA into circulation (Althaher, 2022). NEFA concentrations typically rise around parturition and peak within the first two weeks of lactation, serving as a direct indicator of NEB. Denis-Robichaud (2025) proposed that postpartum plasma NEFA levels exceeding 0.7 mmol/L reflect a severe NEB status. Although moderate NEFA mobilization is physiologically important, providing a substantial source of energy and contributing up to 40% of milk fat in early lactation, excessive NEFA release places a considerable metabolic burden on the liver, leading to triglyceride (TAG) accumulation and increasing the risk of fatty liver and ketosis (Grummer, 1993).

BHBA and acetone are the primary ketone bodies generated in the liver when acetyl-CoA production exceeds the capacity of the TCA cycle (Baird, 1982). At physiological levels, they serve as alternative energy substrates for the brain, cardiac muscle, and skeletal muscle. However, their excessive accumulation in blood reflects a severe NEB state and markedly increases the risk of subclinical or clinical ketosis. Elevated postpartum plasma BHBA and acetone concentrations have been associated with reduced milk yield, impaired reproductive performance, and an increased incidence of secondary metabolic disorders such as displaced abomasum (Suthar et al., 2013; Berge and Vertenten, 2014).

In addition, NEB is accompanied by profound endocrine changes, most notably a reduction in circulating IGF-1. This will decrease hepatic synthesis under energy deficiency. Low IGF-1 concentrations are strongly linked to delayed resumption of ovarian cyclicity, poor follicular development, and reduced conception rates in high-yielding cows (Butler, 2003; Wathes et al., 2007c). These findings underscore IGF-1 as an integrative biomarker linking energy metabolism with reproductive efficiency (Velazquez et al., 2008).

5.3. Milk biomarkers associated with negative energy balance

Compared to blood, milk offers a non-invasive and practical source of biomarkers for NEB monitoring. Milk ketone bodies, such as BHBA and acetone, are directly associated with ketosis and can be routinely predicted using Fourier-transform mid-infrared (FT-MIR) spectroscopy (de Roos et al., 2007; Grelet et al., 2016). Citrate has been identified as an early indicator of NEB, as its concentration reflects shifts in the TCA cycle and overall metabolic status (Bjerre-Harpøth et al., 2012).

Milk fatty acid composition also provides valuable information: short-chain fatty acids (SCFA) and medium-chain fatty acids (MCFA), mainly synthesized de novo in the mammary gland, tend to decrease during NEB, while long-chain fatty acids (LCFA) such as C18:0 and C18:1 cis-9 increase, reflecting enhanced mobilization of adipose reserves (Pires et al., 2022; Churakov et al., 2021). In addition, classical production traits such as MY, fat percentage (FP), and protein percentage (PP) are influenced by NEB. The milk FPR has been widely applied as a practical early-lactation indicator of NEB, showing strong genetic correlations with energy balance (Buttchereit et al., 2011).

5.4. Energy deficiency score derived from multivariate milk spectral patterns

In addition to individual milk biomarkers, multivariate metabolic patterns extracted from MIR spectra provide an integrated view of the physiological adaptations associated with NEB. Using 27 MIR-predicted biomarkers including ketone bodies, citrate, and milk fatty acids. Franceschini et al. (2022) applied an unsupervised

clustering approach and identified five physiologically meaningful metabolic clusters in early-lactation cows. Among these, Cluster 4 represented cows exhibiting the most pronounced NEB-related metabolic profile, characterized by elevated predicted BHBA, acetone, citrate, and LCFA, together with reduced de SCFA and MCFA. These features align closely with the known endocrine and biochemical responses to NEB, such as intensified lipolysis, increased ketogenesis, and shifts in energy substrate utilization.

To facilitate quantitative genetic analyses, the Cluster 4 category was subsequently transformed into a continuous Energy Deficiency Score (EDS) as explained in Chapter 2. This transformation retains the biological specificity of the original cluster while providing the resolution necessary for estimating genetic parameters, examining trait relationships, and performing genome-wide association analyses. As a result, EDS serves as a complementary MIR-derived indicator of NEB, capturing complex metabolic signatures that extend beyond single biomarkers and enabling more refined characterization of energy status in early-lactation dairy cows.

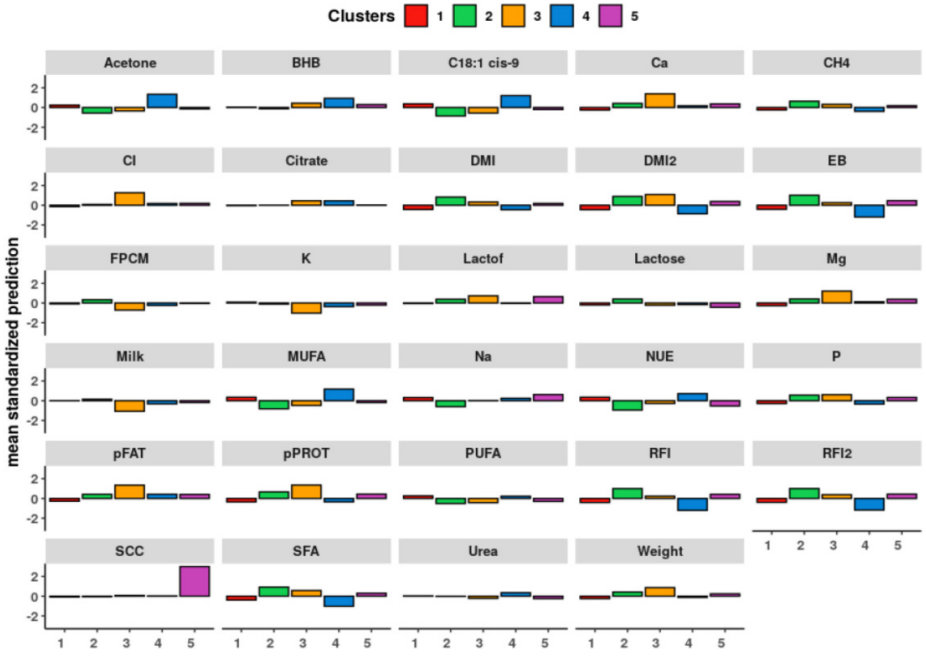


Figure 1-5. Average standardized predicted traits by Cluster 4 (Energy Deficiency Score, EDS) (Franceschini et al., 2022). C18:1 cis-9 = oleic acid; CI = consumption index; EB = energy balance; FPCM = fat-and protein-corrected milk; Lactof = lactoferrin; NUE = nitrogen use efficiency; pPROT = protein content; pFAT = fat content; RFI = residual feed intake (alternate calculation via RFI2); DMI2 = predicted DMI.

6. Genetic analysis methods

6.1. Recursive model for causal inference between traits

In quantitative genetics, multi-trait models (MTM) are commonly used to estimate genetic correlations among traits. While informative, these correlations only describe statistical associations and do not reveal the underlying causal direction. Consequently, it often remains unclear whether one trait causally influences another or whether their association is simply driven by shared genetic or environmental factors. To address this limitation, structural equation models (SEM), and in particular recursive models (RM), have been introduced in animal breeding (Gianola and Sorensen, 2004; Valente et al., 2010; Varona and González-Recio, 2023). RM extend the classical MTM by incorporating structural coefficients (λ) that define unidirectional causal links among traits. This framework allows the decomposition of (co)variance into direct effects (a trait's immediate influence on another) and indirect effects (influences transmitted through intermediate traits).

A practical example relevant to dairy cow physiology is the relationship between NEB and circulating BHB. Physiologically, NEB arises in early lactation when nutrient intake fails to meet energy requirements, leading to mobilization of body fat reserves. This lipid mobilization increases hepatic ketogenesis, thereby elevating BHB concentration. Because NEB affects BHB but BHB does not influence NEB in return, this relationship is inherently directional and can be represented as a simple recursive structure:

$$BHB = \lambda_{21} \cdot NEB + u_{BHB} + e_{BHB},$$

$$NEB = u_{NEB} + e_{NEB}.$$

In this formulation, recursive parameter λ_{21}quantifies the causal effect of NEB on BHB, whereas MTM would only estimate a symmetric genetic correlation between the two traits without distinguishing causation from association. As outlined in Varona and González-Recio (2023) a block LDL decomposition is used on the residual covariance matrix between BHB and NEB to obtain λ_{21}. By explicitly modeling causal pathways, RM enable the separation of direct genetic contributions from mediated responses, yielding a more biologically meaningful interpretation of trait interdependencies particularly for metabolic traits in early lactation dairy cows.

6.2. Genetic correlations estimated from SNP effects

Genetic correlations quantify the degree to which two traits share a common genetic basis (Andersen-Ranberg et al., 2005). Traditionally, these correlations have been estimated using pedigree or marker based relationship matrices within variance component models (e.g., animal models with REML). Such estimates provide global measures of additive genetic covariance between traits, but they do not reveal which

genomic regions contribute to these relationships.

The increasing availability of dense SNP genotypes has enabled the estimation of correlations directly from SNP effects. SNP-based correlation approaches rely on comparing estimated marker effects across traits, often weighted by allele frequency (e.g., 2pq). This allows the decomposition of genetic correlations into genome-wide and chromosome-specific components, providing a finer-scale understanding of the genomic architecture underlying trait associations (Calus and Veerkamp, 2011; Gervais et al. 2017).

In the context of NEB and its related MIR-predicted biomarkers, SNP effect-based correlations are particularly valuable. They allow the identification of shared genomic regions and potential pleiotropic effects, complementing traditional genetic correlation estimates with more detailed insights into the loci contributing to trait interdependence. In this thesis, SNP-based correlations were used to compare MIR-predicted NEB indicators (e.g., LPNEB, LEDS) with metabolic and production traits at both the whole-genome and chromosome levels.

6.3. Application of single-step GWAS to negative energy balance traits in dairy cattle

Single-step genome-wide association studies (ssGWAS) have emerged as a powerful framework for unraveling the genetic architecture of complex traits in livestock. Unlike conventional GWAS, which are restricted to genotyped animals, ssGWAS simultaneously integrates pedigree, phenotype, and genomic information through the unified H-matrix (Legarra et al., 2009; Wang et al., 2012). This approach allows both genotyped and non-genotyped animals to contribute to the analysis, thereby enlarging the effective population size and improving the precision of SNP effect estimation (Cesarani et al., 2022).

In dairy cattle, ssGWAS has been widely adopted to investigate polygenic traits such as production, fertility, health, and metabolic indicators (Klein et al., 2021). Its strength lies in enabling chromosome-level dissection of genetic variance, identification of genomic regions linked to key physiological processes, and quantification of the relative contributions of SNPs to specific traits (Chen et al., 2023).

In the context of this thesis, ssGWAS was applied to LPNEB, LEDS, and multiple MIR-predicted metabolic biomarkers. These traits are particularly challenging to study because NEB is influenced by both genetic and environmental factors and is difficult to measure directly. By leveraging large-scale phenotypic records and dense genomic information, ssGWAS provided novel insights into their genetic architecture, revealed key chromosomal regions associated with lipid mobilization and energy metabolism, and enabled comparisons between MIR-derived composite traits and conventionally measured biomarkers. This framework thus represents a cornerstone

for linking large-scale MIR data to genomic analyses, with direct implications for breeding programs targeting improved resilience to NEB.

7. Research objectives and outline

The primary objective of this thesis is to advance the understanding of NEB in first-parity Holstein cows during early lactation by integrating MIR-derived phenotypes and genomic analyses. This study aims to explore the relationships between LPNEB and key metabolic and production biomarkers, and to investigate the genomic architecture of both LPNEB and LEDS. Collectively, these objectives are intended to establish MIR-based proxies as reliable tools for large-scale phenotyping and to identify genomic regions and candidate genes associated with the regulation of energy balance.

To address these objectives, the thesis is organized into three main studies. The first investigates the phenotypic and genetic correlations between LPNEB, LEDS, 15 biomarkers, and 3 production traits using a 20-trait repeatability model, and further explores the causal effects of 19 studied traits on LPNEB through a recursive model. The second study estimates the genomic correlations between LPNEB and 19 studied traits based on SNP effects, and evaluates the independent contributions of eight biomarkers to LPNEB and LEDS. The third study applies ssGWAS to LPNEB and LEDS to detect genomic regions and candidate genes involved in energy metabolism. Collectively, these studies provide novel insights into the utility of MIR-derived indicators for understanding NEB and its associations with multiple biomarkers, as well as for elucidating their underlying genomic architecture.

7.1. Outline

The thesis consists of seven chapters, with five forming the main content.

In Chapter 1, provides the biological background of energy balance in dairy cows, explains the causes and consequences of NEB, reviews methods and biomarkers used to assess NEB, and introduces the genetic approaches applied in this thesis. the research objectives are also presented.

In Chapter 2, relationships between predicted NEB and biomarkers examines the phenotypic and genetic relationships between PNEB, RNEB, and 19 metabolic and production traits. Genetic parameters, correlations, and causal effects are estimated to understand how these traits reflect early-lactation NEB.

In Chapter 3, genetic architecture of LPNEB, LEDS, and biomarkers compares the genetic architecture of two MIR-derived NEB indicators (LPNEB and LEDS) and several biomarkers using SNP-based genetic correlations and independent trait-contribution analyses across chromosomes.

In Chapter 4, genome-wide association analyses of LPNEB and LEDS.

Identifies genomic regions and candidate genes associated with LPNEB and LEDS using single-step GWAS. Functional annotations are used to highlight biological pathways related to energy metabolism.

In Chapter 5, general discussion, conclusions, and perspectives.

8. References

Althaher, A.R. 2022. An Overview of Hormone-Sensitive Lipase (HSL). The Scientific World Journal, 2022:1964684. https://doi.org/10.1155/2022/1964684.

Andersen-Ranberg, I. M., G. Klemetsdal, B. Heringstad, and T. Steine. 2005. Heritabilities, genetic correlations, and genetic change for female fertility and protein yield in Norwegian dairy cattle. J. Dairy Sci. 88:348–355. https://doi.org/10.3168/jds.S0022-0302(05)72694-1.

Baird, G. D. 1982. Primary ketosis in the high-producing dairy cow: Clinical and subclinical disorders, treatment, prevention, and outlook. J. Dairy Sci. 65:1–10. https://doi.org/10.3168/jds.S0022-0302(82)82146-2.

Banos, G., and M. P. Coffey. 2010. Genetic association between body energy measured throughout lactation and fertility in dairy cattle. Animal 4:189–199. https://doi.org/10.1017/S1751731109991182.

Bastin, C., L. Laloux, A. Gillon, C. Bertozzi, S. Vanderick, and N. Gengler. 2007. First results of body condition score modeling for Walloon Holstein cows. Interbull Bull. 37:170–174.

Berge, A.C., and G. Vertenten. 2014. A field study to determine the prevalence, dairy herd management systems, and fresh cow clinical conditions associated with ketosis in western European dairy herds. J. Dairy Sci. 97:2145–2154. https://doi.org/10.3168/jds.2013-7163.

Bisinotto, R. S., L. F. Greco, E. S. Ribeiro, N. Martinez, F. S. Lima, C. R. Staples, W. W. Thatcher, and J. E. P. Santos. 2012. Influences of nutrition and metabolism on fertility of dairy cows. Anim. Reprod. 9:260–272.

Bjerre-Harpøth, V., N.C. Friggens, V.M. Thorup, T. Larsen, B.M. Damgaard, K.L. Ingvartsen, and K.M. Moyes. 2012. Metabolic and production profiles of dairy cows in response to decreased nutrient density to increase physiological imbalance at different stages of lactation. J. Dairy Sci. 95: 2362-2380. https://doi.org/10.3168/jds.2011-4419.

Brouwer E. 1965. Report of sub-committee on constants and factors. Proceedings of the 3rd Symposium on Energy Metabolism. Troon, Scotland. Academic Press, London, UK. 441–443.

Butler, W. R. 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. Livest. Prod. Sci. 83:211–218. https://doi.org/10.1016/S0301-6226(03)00112-X.

Buttchereit, N., E. Stamer, W. Junge, and G. Thaller. 2011. 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.

Caixeta, L.S., B.O. Omontese. 2021 Monitoring and Improving the Metabolic Health of Dairy Cows during the Transition Period. Animals. 11:352. https://doi.org/10.3390/ani11020352.

Calus, M. P. L., and R. F. Veerkamp. 2011. Accuracy of multi-trait genomic selection using different methods. Genet. Sel. Evol. 43:26. https://doi.org/10.1186/1297-9686-43-26.

Casaro, S., J. Pérez-Báez, R. S. Bisinotto, R. C. Chebel, J. G. Prim, T. D. Gonzalez, G. Carvalho Gomes, S. Tao, I. M. Toledo, B. C. do Amaral, J. M. Bollati, M. G. Zenobi, N. Martinez, G. E. Dahl, J. E. P. Santos, and K. N. Galvão. 2024. Association between prepartum body condition score and prepartum and postpartum dry matter intake and energy balance in multiparous Holstein cows. J. Dairy Sci. 107:4381–4393. https://doi.org/10.3168/jds.2023-24047.

Cesarani, A., D. Lourenco, S. Tsuruta, A. Legarra, E. L. Nicolazzi, P. M. VanRaden, and I. Misztal. 2022. Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor. J. Dairy Sci. 105:5141–5152. https://doi.org/10.3168/jds.2021-21505.

Chen, Y., H. Atashi, C. Grelet, R. R. Mota, S. Vanderick, H. Hu, GplusE Consortium, and N. Gengler. 2023. Genome-wide association study and functional annotation analyses for nitrogen efficiency index and its composition traits in dairy cattle. J. Dairy Sci. 106:3397–3410. https://doi.org/10.3168/jds.2022-22351.

Churakov, M., J. Karlsson, A. E. 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.

de Roos, A.P.W., H.J.C.M. van den Bijgaart, J. Hørlyk, G. de Jong. 2007. Screening for Subclinical Ketosis in Dairy Cattle by Fourier Transform Infrared Spectrometry. J. Dairy Sci. 90: 1761-1766. https://doi.org/10.3168/jds.2006-203.

Denis-Robichaud, J., I. Nicola, H. Chupin, J.-P. Roy, S. Buczinski, V. Fauteux, N. Picard-Hagen, and J. Dubuc. 2025. Nonesterified fatty acids during the dry period and their association with peripartum disorders, culling, and pregnancy in dairy cows. JDS Communications. 6:688-693. https://doi.org/10.3168/jdsc.2025-0784.

Dezetter, C., F. Bidan, L. Delaby, F. Blanc, S. Freret, and N. Bedere. 2024. Association between body condition profiles, milk production, and reproduction performance in Holstein and Normande cows. J. Dairy Sci. 107:11621–11638. https://doi.org/10.3168/jds.2024-24766.

Dietz, J., J. Schwartz. 1991. Growth hormone alters lipolysis and hormone-sensitive lipase activity in 3T3-F442A adipocytes. Metabolism. 40:800-806.https://doi.org/10.1016/0026-0495(91)90006-I.

Drackley, J. K. 1999. Biology of dairy cows during the transition period: The final frontier? J. Dairy Sci. 82:2259–2273. https://doi.org/10.3168/jds.S0022-0302(99)75474-3.

Ducháček, J., L. Stádník, M. Ptáček, J. Beran, M. Okrouhlá, and M. Gašparík. 2020.. Negative Energy Balance Influences Nutritional Quality of Milk from Czech Fleckvieh Cows due Changes in Proportion of Fatty Acids. Animals. 10: 563. https://doi.org/10.3390/ani10040563.

Duffield, T. F., K. D. Lissemore, B. W. McBride, and K. E. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92:571–580. https://doi.org/10.3168/jds.2008-1507.

Elischer, M. F., M. E. Arceo, E. L. Karcher, and J. M. Siegford. 2013. Validating the accuracy of activity and rumination monitor data from dairy cows housed in a pasture-based automatic milking system. J. Dairy Sci. 96:6412–6422. https://doi.org/10.3168/jds.2013-6790.

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. Animal Reproduction Science 144:60–71. https://doi.org/10.1016/j.anireprosci.2013.11.007.

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.

Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M.-L. Vanrobays, A. Lainé, S. Vanderick, C. Grelet, A. Vanlierde, E. Froidmont, P. Dardenne. 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. J. Dairy Sci. 99:4071-4079. https://doi.org/10.3168/jds.2015-10140.

Gervais, O., R. Pong-Wong, P. Navarro, C. S. Haley, and Y. Nagamine. 2017. Antagonistic genetic correlations for milking traits within the genome of dairy cattle. PLoS One 12:e0175105. https://doi.org/10.1371/journal.pone.0175105.

Gianola, D., and D. Sorensen. 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. Genetics 167:1407–1424. https://doi.org/10.1534/genetics.103.025734.

Gillund, P., O. Reksen, Y.T. Gröhn, and K. Karlberg. 2001. Body Condition Related to Ketosis and Reproductive Performance in Norwegian Dairy Cows. J. Dairy Sci. 84:1390-1396. https://doi.org/10.3168/jds.S0022-0302(01)70170-1.

Grelet, C. P. Dardenne, H. Soyeurt, J.A. Fernandez, A. Vanlierde, F. Stevens, N. Gengler, and F. Dehareng. 2021. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. Methods. 186:97-111.https://doi.org/10.1016/j.ymeth.2020.07.012.

Grelet, C., C. Bastin, M. Gelé, J. B. Davière, M. Johan, A. Werner, R. Reding, J. A. Fernandez Pierna, F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825. https://doi.org/10.3168/jds.2015-10477.

Grummer, R. R. 1993. Etiology of lipid-related metabolic disorders in periparturient dairy cows. J. Dairy Sci. 76:3882–3896. https://doi.org/10.3168/jds.S0022-0302(93)77729-2.

Harder, I., E. Stamer, W. Junge, and G. Thaller. 2019. Lactation curves and model evaluation for feed intake and energy balance in dairy cows. J. Dairy Sci. 102:7204–7216. https://doi.org/10.3168/jds.2018-15300.

Huhtanen, P., and A.R. Bayat. 2025. Potential of novel feed efficiency traits for dairy cows based on respiration gas exchanges measured by respiration chambers or GreenFeed. J. Dairy Sci. 108:12340–12351. https://doi.org/10.3168/jds.2025-26683.

Ingvartsen, K. L., and K. Moyes. 2013. Nutrition, immune function, and health of dairy cattle. Animal 7:112–122. https://doi.org/10.1017/S175173111200170X.

Kaniamuthan, S., A. Manimaran, A. Kumaresan, P. R. Wankhade, T. Karuthadurai, M. Sivaram, and D. Rajendran. 2025. Biochemical indicators of energy balance in blood and other secretions of dairy cattle: A review. Agricultural Reviews 46:247–255. https://doi.org/10.18805/ag.R-2571.

Klein, S.-L., T. Yin, H. H. Swalve, and S. König. 2021. Single-step genomic best linear unbiased predictor genetic parameter estimations and genome-wide associations for milk fatty acid profiles, interval from calving to first insemination, and ketosis in Holstein cattle. J. Dairy Sci. 104:10921–10933. https://doi.org/10.3168/jds.2021-20416.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663. https://doi.org/10.3168/jds.2009-2061.

Leroy, J. L. M. R., T. Vanholder, G. Opsomer, A. Van Soom, and A. de Kruif. 2006. The in vitro development of bovine oocytes after maturation in glucose and β-hydroxybutyrate concentrations associated with negative energy balance in dairy cows. Reproduction in Domestic Animals 41:119–123. https://doi.org/10.1111/j.1439-0531.2006.00650.x.

McArt, J. A. A., D. V. Nydam, and G. R. Oetzel. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. J. Dairy Sci. 95:5056–5066. https://doi.org/10.3168/jds.2012-5443.

McParland, S., E. Kennedy, E. Lewis, S. G. Moore, B. McCarthy, M. O'Donovan, and D. P. Berry. 2015. Genetic parameters of dairy cow energy intake and body energy status predicted using mid-infrared spectrometry of milk. J. Dairy Sci. 98:1310–1320. https://doi.org/10.3168/jds.2014-8892.

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.

Mellouk, N., C. Rame, D. Naquin, Y. Jaszczyszyn, J.-L. Touzé, E. Briant, D. Guillaume, T. Ntallaris, P. Humblot, and J. Dupont. 2019. Impact of the severity of negative energy balance on gene expression in the subcutaneous adipose tissue of periparturient primiparous Holstein dairy cows: Identification of potential novel metabolic signals for the reproductive system. PLoS One 14:e0222954. https://doi.org/10.1371/journal.pone.0222954.

Mulvany, P.M. 1977. Dairy cow condition scoring. Handout No. 4468. Natl. Inst. Res. Dairying 4:349–353.

NRC. 2001. Nutrient Requirements of Dairy Cattle. 7th edition. National Academy Press, Washington, DC, USA.

Pires, J. A. A., T. Larsen, and C. Leroux. 2022. Milk metabolites and fatty acids as noninvasive biomarkers of metabolic status and energy balance in early-lactation cows. J. Dairy Sci. 105:201–220. https://doi.org/10.3168/jds.2021-20465.

Roche, J.R. K.A. Macdonald, C.R. Burke, J.M. Lee, and D.P. Berry. 2007. Associations Among Body Condition Score, Body Weight, and Reproductive Performance in Seasonal-Calving Dairy Cattle. J. Dairy Sci. 90:376-391. https://doi.org/10.3168/jds.S0022-0302(07)72639-5.

Roche, J.R., N.C. Friggens, J.K. Kay, M.W. Fisher, K.J. Stafford, D.P. Berry. 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. J. Dairy Sci. 92:5769-5801. https://doi.org/10.3168/jds.2009-2431.

Singh, A. K., and C. Bhakat. 2022. The relationship between body condition score and milk production, udder health and reduced negative energy balance during initial lactation period: A review. Iran. J. Appl. Anim. Sci. 12:1–10.

Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695. https://doi.org/10.3168/jds.S0022-0302(06)72409-2.

Sundrum, A. 2015. Metabolic Disorders in the Transition Period Indicate that the Dairy Cows' Ability to Adapt is Overstressed. Animals, 5:978-1020. https://doi.org/10.3390/ani5040395.

Suriyasathaporn, W., C. Heuer, E. N. Noordhuizen-Stassen, and Y. H. Schukken. 2000. Hyperketonemia and the impairment of udder defense: A review. Vet. Res. 31:397–412. https://doi.org/10.1051/vetres:2000128.

Suthar, V. S., J. Canelas-Raposo, A. Deniz, and W. Heuwieser. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. J. Dairy Sci. 96:2925–2938. https://doi.org/10.3168/jds.2012-6035.

Thorup, V. M., D. Edwards, and N. C. Friggens. 2012. On-farm estimation of energy balance in dairy cows using only frequent body weight measurements and body condition score. J. Dairy Sci. 95:1784–1793. https://doi.org/10.3168/jds.2011-4631.

Valente, B. D., R. L. Rosa, D. Gianola, G. J. M. Rosa, and K. A. Weigel. 2010. Searching for recursive causal structures in multivariate quantitative genetics mixed models. Genetics 185:633–644. https://doi.org/10.1534/genetics.109.112979.

Varona, L., and O. González-Recio. 2023. Invited review: Recursive models in animal breeding: Interpretation, limitations, and extensions. J. Dairy Sci. 106: 2198-2212. https://doi.org/10.3168/jds.2022-22578.

Velazquez, M. A., L. J. Spicer, and D. C. Wathes. 2008. The role of endocrine insulin-like growth factor-I (IGF-I) in female bovine reproduction. Domest. Anim. Endocrinol. 35:325–342. https://doi.org/10.1016/j.domaniend.2008.07.002.

Walsh, S. W., E. J. Williams, and A. C. O. Evans. 2011. A review of the causes of poor fertility in high milk producing dairy cows. Anim. Reprod. Sci. 123:127–138. https://doi.org/10.1016/j.anireprosci.2010.12.001.

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.

Wang, Z., Y. Song, C. Zhao, Y. Bai, F. Zhang, C. Xia, S. Fu, H. Zhang, C. Xu, and L. Wu. 2022. The relationship of negative energy balance (NEB) and energy metabolism, milk production and reproductive performance during early lactation in dairy cows in Heilongjiang, China. Vet. arhiv 92:223–232. https://doi.org/10.24099/vet.arhiv.1377.

Wathes, D. C., M. Fenwick, Z. Cheng, N. Bourne, S. Llewellyn, D. G. Morris, D. Kenny, J. Murphy, and R. Fitzpatrick. 2007a. 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.

Wathes, D. C., Z. Cheng, N. Bourne, V. J. Taylor, M. P. Coffey, and S. Brotherstone. 2007c. Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield and body condition score in the periparturient period. Domest. Anim. Endocrinol. 33:203–225. https://doi.org/10.1016/j.domaniend.2006.05.004.

Wathes, D.C., N. Bourne, Z. Cheng, G.E. Mann, V.J. Taylor, and M.P. Coffey. 2007b. Multiple Correlation Analyses of Metabolic and Endocrine Profiles with Fertility in Primiparous and Multiparous Cows. J. Dairy Sci. 90:1310-1325. https://doi.org/10.3168/jds.S0022-0302(07)71619-3.

White, H. M. 2015. The role of TCA cycle anaplerosis in ketosis and fatty liver in periparturient dairy cows. Animals 5:793–802. https://doi.org/10.3390/ani5030384.

Wildman, E. E., G. M. Jones, P. E. Wagner, R. L. Boman, H. F. Troutt Jr., and T. N. Lesch. 1982. A dairy cow body condition scoring system and its relationship to selected production characteristics. J. Dairy Sci. 65:495–501. https://doi.org/10.3168/jds.S0022-0302(82)82223-6.

Xu, W., J. Vervoort, E. Saccenti, B. Kemp, R. J. van Hoeij, and A. T. van Knegsel. 2020. Relationship between energy balance and metabolic profiles in plasma and milk of dairy cows in early lactation. J. Dairy Sci. 103:4795–4805. https://doi.org/10.3168/jds.2019-17777.

Zachut, M., M. Šperanda, A.M. De Almeida, G. Gabai, A. Mobasheri, 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.

2

Chapter II Exploring the relationship between predicted negative energy balance and its biomarkers of Holstein cows in first-parity early lactation

Adapted from: 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.

Foreword

Negative energy balance (NEB) in dairy cows adversely affects their health, fertility, and milk production. However, direct measurement of NEB is costly and impractical for large-scale application. To address this, various biomarkers have been developed as indirect indicators of NEB. Despite their potential, the genetic relationships between NEB and these biomarkers remain poorly understood.

This chapter investigates the genetic associations between predicted NEB, a novel Energy Deficit Score (EDS), fifteen candidate NEB-related biomarkers, and three production traits. The results indicate that eight traits such as blood non-esterified fatty acids (NEFA) exhibited strong genetic correlations with predicted NEB, highlighting their relevance for future selection strategies.

Abstract

The negative energy balance (NEB) state in dairy cows is a critical factor affecting health, reproduction, and production, particularly during early lactation. Multiple blood and milk biomarkers change when dairy cows are in the NEB state. Direct measurement of NEB is impractical for large-scale use due to costs, necessitating reliance on indirect predictors such as milk mid-infrared (MIR) spectrometry based predicted biomarkers. However, the genetic relationships between NEB and its potential biomarkers remain unclear. This study aimed to (1) compare measured reference NEB (RNEB) with MIR-predicted NEB (PNEB), a novel Energy Deficit Score (EDS), 15 biomarkers, and 3 production traits; (2) estimate genetic parameters among these traits using a 20-trait repeatability model, quantifying the ability of the 19 other studied traits (logit-transformed EDS (LEDS), 15 biomarkers, and three production traits) to genetically predict logit-transformed PNEB (LPNEB); and (3) evaluate the causal effects of LPNEB on the 19 traits through a recursive model. Two datasets were utilized: Dataset I (127 cows, 965 records) provided reference data for objective (1), and dataset II (25,287 first-parity cows, 30,634 records) enabled genetic analysis used for objectives (2) and (3). Traits were analyzed using Pearson correlations, multiple-diagonalization EM-REML based genetic parameter estimation, and recursive modeling. The studied traits had moderate to moderate-high h2 ranging from 0.16 to 0.38. The genetic correlations between LPNEB and the studied traits ranged from -0.60 to 0.85 for LEDS and 0.87 for MIR-predicted blood non-esterified fatty acids (NEFA). Analysis of genetic predictability of LPNEB revealed that the 19 other traits together explained 89% of the genetic variance of LPNEB, with all 15 biomarkers alone contributing the largest fraction with 82%, LEDS alone 65%, NEFA alone 62%, and all traits except LEDS 85%, indicating that LEDS contains useful additional information. Recursive modeling further identified 8 traits, including NEFA and LEDS, as highly dependent on LPNEB, highlighting their potential as robust biomarkers. This study demonstrates the utility of MIR-predicted traits for understanding the genetic mechanisms of NEB and its potential for integration into breeding programs, while emphasizing cautious interpretation of these results due to limitations of MIR-predictions of studied traits to represent directly measured traits.

Key words: heritability, milk mid-infrared, genetic correlation, phenotypic correlation, recursive model

1. Introduction

Energy balance (EB) in dairy cows refers to the difference between the energy intake and the energy expended for maintenance and production activities, such as milk production, growth, and reproduction (Baumgard et al., 2006). In the early lactation stage, when the energy intake cannot meet the demands of milk production, most dairy cows are in the state of negative EB (NEB) (Churakov et al., 2021). For example, a recent study considered that 75.2% of cows in the UK were in NEB during the first 20 days in milk (DIM) (Macrae et al., 2019). Other studies have shown that the extent and duration of the postpartum NEB state is an important factor affecting the reproduction and health of dairy cows (Collard et al., 2000). If a dairy cow stays in the NEB state for a long period, the risk of metabolic diseases (e.g., ketosis) will increase. In addition, the NEB state is associated with decreased conception rates, increased early embryo mortality, and increased estrus silencing in dairy cows (Zachut et al., 2020).

Monitoring the occurrence of NEB state in dairy cows in early lactation is essential for optimizing the health condition of dairy herds (Churakov et al., 2021) and generating data for breeding purposes for reduced NEB risk. Although metabolic chambers are the gold standard method for detecting energy status, this equipment is too expensive for large-scaled applications (Churakov et al., 2021). Therefore, there are a few reports with very small number of records on the genetic analysis of measured EB (e.g., Berry et al., 2007; Liinamo et al., 2012). Moreover, body condition score (BCS) has traditionally been used to monitor animals and identify those which are in NEB (Stádník et al., 2002). However, human assessed BCS scores are subjective, labor-intensive, and not easily available as repeated records in early lactation. Also, even if precision livestock farming tools based on computer vision are becoming available, these technologies remain expensive and are not easy to implement in commercial farms. Therefore, the development of an objective, accurate, quick, inexpensive, and easy procedure to detect NEB in dairy farms is still needed.

For dairy cows, being in a NEB state is a complex physiological metabolic process involving multiple changes in the metabolic pathways and energy sources, modifying blood content of non-esterified fatty acid (NEFA), beta-hydroxybutyrate acid (BHBA) or Insulin-like growth factor 1 (IGF-1) (Zachut et al., 2020; Xu et al., 2020; Pires et al., 2022). Therefore, when testing NEB on farms, veterinarians traditionally rely on directly measured levels of these biomarkers (Macrae et al., 2019). However, direct measuring of these biomarkers is always based on blood samples that are obtained in an invasive manner, which is not recommended on a large scale due to animal welfare reasons. Furthermore, for practical reasons, most studies used only 1 or 2 biomarkers, such as NEFA, BHBA, for predicting NEB state. These two issues could potentially be addressed using milk mid-infrared (MIR) spectrometry-based predictions. Indeed,

several suitable predictors using milk MIR spectrometry, which is obviously a non-invasive method, for milk but also blood biomarkers have been reported (Grelet et al., 2021). However, some biomarkers remain potentially poorly predicted by MIR. Therefore, recently, Franceschini et al. (2022) used an unsupervised machine learning method to integrate multiple biomarkers (mostly MIR-predicted biomarkers) into a clustering approach to develop a novel trait called hereafter the Energy Deficit Score (EDS).

Mid-infrared spectrometric data are used to predict several biomarkers and can be considered as an efficient approach to predict EB in dairy cows as well. However, as already stated, the coefficients of determination of the different prediction models vary across studies (from 0.48 to 0.78) (McParland and Berry, 2016; Smith et al., 2019; Ho et al., 2020) indicating that not all biomarkers can be predicted in a reliable manner using MIR spectrometry. McParland et al. (2015) and Smith et al. (2019) reported genetic parameters of MIR-predicted EB with larger datasets. However, in all these studies, the genetic relationships between NEB state and multiple biomarkers and economically relevant traits (e.g., production traits such as fat percentage (FP), protein percentage (PP), and milk yield (MY)) remain unclear.

The objectives of this study were therefore to (1) compare measured reference NEB (RNEB) to MIR-predicted NEB (PNEB), Energy Deficit Score (EDS), 15 biomarkers, and 3 production traits; (2) estimate the genetic parameters among the mentioned traits using a 20-trait repeatability model allowing the quantification of the ability of the 19 other traits (i.e., biomarkers) to genetically predict the logit transformed predicted NEB (LPNEB); (3) use these genetic parameters to estimate the causal effects of LPNEB on the 19 mentioned traits through a recursive model. Achieving these objectives will allow to highlight the role of other traits (i.e., biomarkers) in predicting NEB, as well as the causal effects of NEB on other traits. However, due to the reliance on MIR-predicted traits in this study, conclusions will be drawn cautiously to avoid overinterpretation.

2. Materials and methods

The study franework is shown in Figure 2-1.

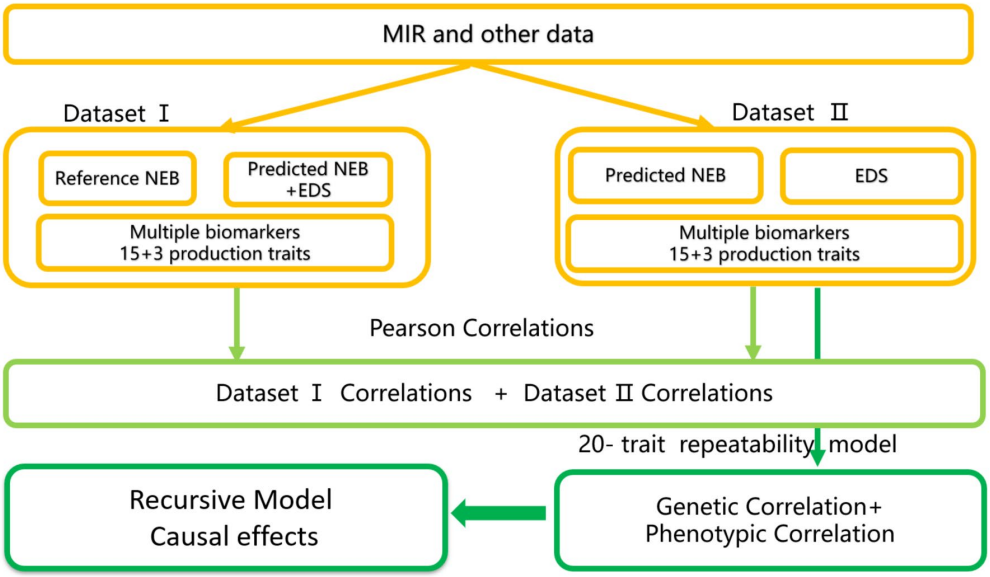


Figure 2-1. Workflow illustrating the data structure and analytical procedures used in this study.

2.1. Data collection and editing

Because this study used many MIR-predicted biomarkers in a first step, we used a dataset I, which included RNEB, PNEB, and 19 other traits. Dataset I allowed to compare observed correlations of RNEB with PNEB and the other 19 traits. It was collected by the GplusE Project (https://gpluse.eaap.org/) on 3 experimental farms located in 3 different countries (Agri-Food and Biosciences Institute, Belfast, UK; Aarhus University, Aarhus, Denmark; and University College Dublin, Dublin, Ireland). This dataset provided access to 1,033 directly measured reference EB (REB) records on 129 dairy cows collected from DIM 5 to 50. The diet of the animals and the analysis methods have been previously reported by Grelet et al. (2020). In a second part of the study, a larger dataset (including 20 traits) was used to do genetic analysis, hereafter referred to as dataset II. We used dataset II to calculate the genetic relationships among 20 traits, revealing their genetic components and interrelationships. Dataset II was compiled during the official recording of milk data in the Walloon Region of Belgium from 2012 to 2019. Both datasets encompassed the same 20 traits, except for REB, which was exclusively included in dataset I. Please note that to distinguish directly measured biomarkers from those predicted by MIR

spectrometry, abbreviations will be employed solely for the MIR-predicted biomarkers. The considered traits (n = 20) in both datasets include predicted EB (PEB), EDS, 15 biomarkers, and 3 production traits (FP, PP, MY). The choice of the 15 biomarkers was based on their recognized involvement in energy metabolism. The PEB and the selected biomarkers (n = 15) were predicted by milk MIR spectrometry or MIR spectra data plus MY and parity. All the used prediction equations were described in previous publications as outlined in Table 2-1. Milk MIR spectra, FP, and PP were generated by commercial instruments FT2, FT6000 spectrometers (Foss Analytics) and on a Standard Lactoscope FT-MIR automatic (Delta Instruments). MIR-infrared spectral data of the 2 datasets were standardized by the method proposed by Grelet et al. (2015). Milk yield was recorded during milk sampling. As described by Franceschini et al. (2022), the EDS trait was defined using the agglomerative hierarchical clustering method based on 27 MIR-based predictors. However, new observations were predicted directly using MIR spectral data. More details on the methodology and the 27 MIR spectra-based predictors can be found in Franceschini et al. (2022).

Table 2-1. Description of the 16 traits[1] prediction equations used in this study.

Traits	Unit	R^2cv[2]	RMSEcv[3]	Reference
EB	Mcal/d	0.43	N/A[4]	Grelet et al., 2017
In blood				
B_BHBA	mmol/L	0.70	0.27	Grelet et al., 2019
NEFA	µekv/L	0.39	344.20	Grelet et al., 2019
IGF-1	mg/L	0.61	44.40	Grelet et al., 2019
GLU	mmol/L	0.44	0.36	Grelet et al., 2019
In milk				
M_BHBA	mmol/L	0.71	0.11	Grelet et al., 2016
CIT	mmol/L	0.90	0.70	Grelet et al., 2016
ACE	mmol/L	0.73	0.25	Grelet et al., 2016
C10:0	g/dL milk	0.91	0.09	Soyeurt et al., 2006
C14:0	g/dL milk	0.93	0.07	Soyeurt et al., 2006
C16:0	g/dL milk	0.94	0.08	Soyeurt et al., 2006
C18:0	g/dL milk	0.84	0.14	Soyeurt et al., 2006
C18:1 cis-9	g/dL milk	0.95	0.08	Soyeurt et al., 2011
SCFA	g/dL milk	0.93	0.07	Soyeurt et al., 2011
MCFA	g/dL milk	0.97	0.05	Soyeurt et al., 2011
LCFA	g/dL milk	0.95	0.07	Soyeurt et al., 2011

[1]EB = energy balance; B_BHBA = blood beta-hydroxybutyrate acid; NEFA = blood non esterified fatty acids; IGF-1= blood insulin-like growth factor 1; GLU = blood glucose; M_BHBA = milk beta-hydroxybutyrate acid; CIT = milk citrate; ACE = milk acetone; C10:0= milk decanoic acid; C14:0 = milk myristic acid; C16:0 = milk palmitic acid; C18:0 = milk stearic acid; C18:1 cis-9 = milk oleic acid; SCFA = milk short-chain fatty acids; MCFA = milk medium chain fatty acids; LCFA = milk long-chain fatty acids.
[2]R^2cv= coefficient of determination of cross-validation.
[3]RMSEcv= root mean square error of cross-validation.
[4]N/A= not applicable because data not available.

Trait transformations. Because PEB does not reflect the relative risk of being in an NEB predicted from MIR spectral data (PNEB), PEB was converted into PNEB by the following formula:

$$\mathrm{PNEB} = 1 - \frac{\mathrm{PEB} - \mathrm{PEB}_{\mathrm{minimum}}}{\mathrm{PEB}_{\mathrm{maximun}} - \mathrm{PEB}_{\mathrm{minimum}}} \quad (2.1)$$

Expressing NEB as a relative value in the observed range of PEB values in a given dataset. The larger the value of PNEB, the worse the energy status of the cow expressed on this relative scale. The probabilities of PNEB and EDS were subjected to a logit-like transformation due to their nonnormal distribution because they were constrained within the interval of 0 to 1. The formula employed for calculating PNEB is expressed as follows: LPNEB = log10(PNEB/(1-PNEB)). To ensure that the values of PNEB remained within a specified range, those below 0.001 were adjusted to 0.001, and those exceeding 0.999 were set to 0.999. This adjustment effectively confines the range of LPNEB to between -3 and +3. The same procedure was used for measured REB generating logit transformed measured reference NEB (LRNEB). These transformations notably enhanced the theoretical properties of LPNEB and LRNEB compared to those of PEB and measured REB, by expressing better relative risk under the specific production circumstances of each dataset. However, this improvement came at the cost of losing the ability of direct comparisons of descriptive parameters, particularly the means of these traits, across variables and datasets. An equivalent formula was then used to compute and to define a novel logit-like transformed energy deficit score (LEDS): LEDS=log10(EDS/(1-EDS)). These transformations are related to odds as a way of expressing the likelihood of an event (here risk of NEB state) occurring compared with it is not occurring. Using base 10 facilitated the treatment of extreme boundaries, here to be set between -3 and +3, and does not affect results. Therefore, in further text the used transformation will be called logit-transformation.

The visual inspection showed MIR-based predictions of blood beta-hydroxybutyrate acid (B_BHBA), insulin-like growth factor 1 (IGF-1), milk beta-hydroxybutyrate acid (M_BHBA), and acetone (ACE) to deviate from a normal distribution (Figure 2-2). Therefore, Log10 transformations were applied (LB_BHBA, LIGF-1, LM_BHBA, LACE) leading to distributions that were closer to normal (Figure 2-2). As a reminder, we are using abbreviations of biomarker names only for the MIR-based predictions.

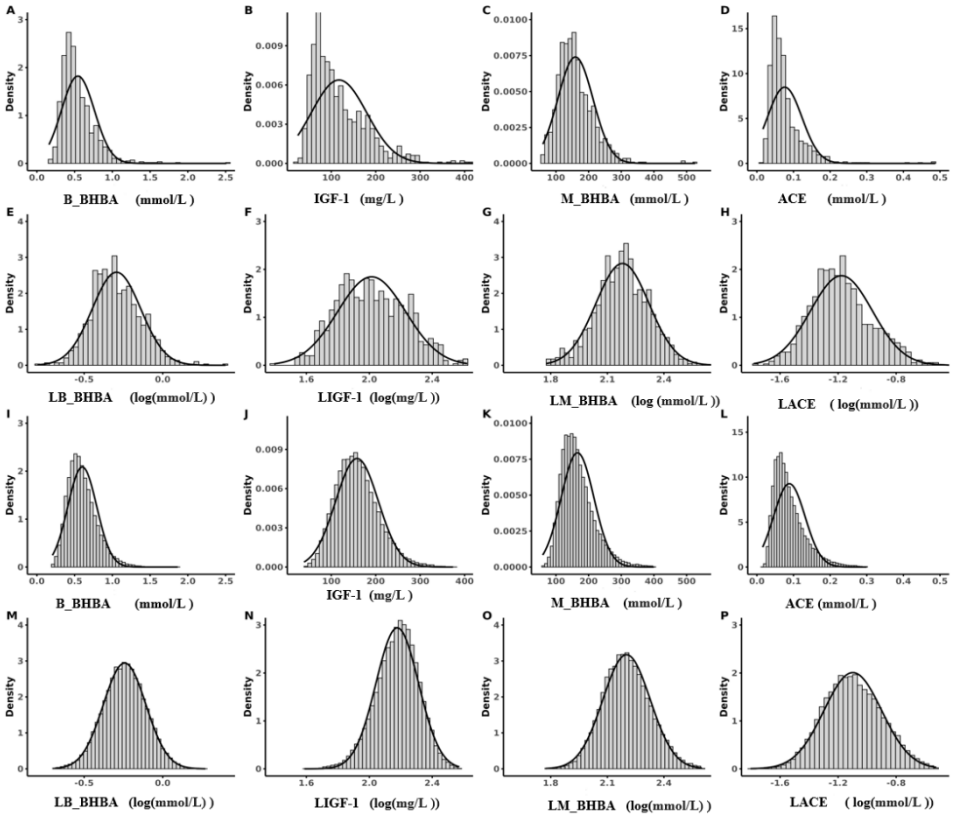


Figure 2-2. Distribution of the standardized 16 traits in dataset I (A–H) and dataset II (I–P). (A) Blood β-hydroxybutyric acid predicted by mid-infrared spectra; (B) blood IGF-1 predicted by mid-infrared spectra; (C) milk β-hydroxybutyric acid predicted by mid-infrared spectra; (D) milk acetone predicted by mid-infrared spectra; (E) log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; (F) log10 blood IGF-1 predicted by mid-infrared spectra; (G) log10 milk β-hydroxybutyric acid predicted by mid-infrared spectra; (H) log10 milk acetone predicted by mid-infrared spectra; (I) blood β-hydroxybutyric acid predicted by mid-infrared spectra; (J) blood IGF-1 predicted by mid-infrared spectra; (K) milk β-hydroxybutyric acid predicted by mid-infrared spectra; (L) milk acetone predicted by mid-infrared spectra; (M) log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; (N) log10 blood IGF-1 predicted by mid-infrared spectra; (O) log10 milk β-hydroxybutyric acid predicted by mid-infrared spectra; (P) log10 milk acetone predicted by mid-infrared spectra.

Data editing. All the following procedures were applied to both datasets. Milk yield, FP, and PP were limited from 3 to 99 kg/day, 1 to 9 %, and 1 to 7%, respectively (ICAR, 2022). The filter proposed by Chen et al. (2021) was used for cleaning the selected milk MIR-predicted traits (PNEB and 15 biomarkers in blood and milk). The Global H distances (Whitfield et al., 1987) were computed for each standardized MIR spectrum in this study in relation to the MIR spectra from the calibration dataset

corresponding to various traits. Only those MIR records with a Global H value of 3 or less were retained. Furthermore, all predicted traits were constrained within a DIM range of 5 to 50 because the predictive models for certain biomarkers are exclusively applicable within this DIM range (e.g., PEB, B_BHBA).

Used datasets and traits. A total of 965 data records on 127 Holstein cows (with parity levels varying from 1 to 7, mean: 2.75, SD: 1.54) across three herds were retained for dataset I. Additionally, Dataset II comprised 30,634 records from 25,287 first parity Holstein cows distributed in 508 herds. The pedigrees for the animals included in dataset II were sourced from the Walloon genetic evaluation system, which incorporated four generations of ancestors, culminating in a total of 74,662 animals, including 4,700 bulls.

The studied traits were LRNEB (only dataset I), LPNEB, LEDS, 15 biomarkers including LB_BHBA, NEFA, LIGF-1, glucose (GLU), LM_BHBA, citrate (CIT), LACE, decanoic acid (C10:0), myristic acid (C14:0), palmitic acid (C16:0), stearic acid (C18:0), short-chain fatty acids (SCFA), medium chain fatty acids (MCFA), long-chain fatty acids (LCFA), oleic acid (C18:1 cis-9), and 3 production traits (FP, PP, MY).

2.2. Comparison of measured EB and predicted EB with 19 other traits

To examine the associations between LPNEB (LRNEB, for dataset I) and the remaining 19 traits, Pearson correlations (r) and their corresponding *P*-values were computed for both datasets. This approach was considered essential because dataset I was too limited in size to support a robust genetic analysis.

To eliminate the influence of parity on traits in dataset I, we applied a linear model to adjust traits, standardizing them to the baseline level of parity 1. The specific approach is as follows. First inside a given trait a linear model for each observation i inside parity class ($y_{ij} = \beta_j + \varepsilon_{ij}$) was constructed based on the parity classification variable, and the solution corresponding to parity 1, here β_1 was extracted. The adjusted trait value for y_{ij} ($y_{ij}^{adjusted}$) was then calculated using the formula: $y_{ij}^{adjusted} = y_{ij} - \beta_j + \beta_1$, where β_j represents the predicted value from the model, and β_1 represents the effect of parity 1. The adjusted trait values, labeled as "adjusted traits" are corrected to the level of parity 1, ensuring theoretical consistency when comparing traits across different parities.

2.3. Genetic parameters estimation

(Co)variance component estimation. The estimation of (co)variance components for 20 traits within dataset II was conducted utilizing a 20-trait repeatability model. This approach was facilitated by the application of canonical transformation across

multiple matrices, employing the multiple diagonalization method as described by Misztal and Lawlor (1995) using the following model:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Mm} + \mathbf{Cc} + \mathbf{Dd} + \mathbf{Qq} + \mathbf{Wp} + \mathbf{Za} + \mathbf{e}, \quad (2.2)$$

where **y** is the vector of observations for the studied traits (n = 20: 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 to 24 months, 25, 26, 27, 28, 29, 30, 31 to 32 months, 33 and 34 months, and 35 to 37 months); **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 to 6, DIM 7 to 8, DIM 9 to 10, and DIM 11 to 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; **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 and goodness of fit obtained visualized comparing mean values of traits with sums of class and regression effects at each DIM.

The expected values and variances associated with this model were defined as follows:

$$E(\mathbf{y}) = \mathbf{Hh} + \mathbf{Mm} + \mathbf{Cc} + \mathbf{Dd} + \mathbf{Qq}, \quad (2.3)$$

$$E(\mathbf{p}) = E(\mathbf{a}) = E(\mathbf{e}) = 0. \quad (2.4)$$

The matrices Var(**p**) and Var(**a**) contained 20×20 basic (co)variance blocks among the 20 traits. For Var(*e*), the off-diagonal elements were modeled as non-zero values, representing residual covariance between correlated traits. This approach ensures that the environmental covariance among traits is not concentrated on the permanent environmental effects.

Due to the informative pedigree, genetic and permanent environmental effects could be separated. Given the fact that we needed reliable (co)variance components for 20 traits, a fast and robust expectation maximization REML (EM-REML) algorithm was used that integrated the multiple-diagonalization algorithm as implemented by Misztal et al. (1995) in the MTC program (http://nce.ads.uga.edu/~ignacy/numpub/mtc/mtcman). This MTC program was shown to be adapted for situations where many (co)variances had to be estimated at the same time (e.g., Rustin et al., 2009) using deceleration during (co)variance components updating as first proposed by Wiggans et al. (2006). Computing time for a round of EM-REML estimation was approximately 20 times faster when using multiple-diagonalization. Relative off-diagonals at convergence, as presented by Misztal and Lawlor (1995), were available to check the efficiency of the multiple-

diagonalization algorithm. Perfect multiple-diagonalization will result in zero off-diagonal elements. As in this algorithm all transformed matrices need to become diagonal, the values of 0.42×10^{-3} for the permanent environment and 0.11×10^{-3} for genetic (co)variances indicate excellent diagonalization. This low relative off-diagonal elements demonstrate that the optimization process have been effectively converged leading to correct estimations of variances and covariances partitioning variance components.

Heritability, genetic and phenotypic correlations. For each trait (including LPNEB, LEDS, 15 biomarkers, and 3 production traits), the heritability (h^2) was calculated using the following equation:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2} \qquad (2.5)$$

where σ_a^2, σ_p^2 and σ_e^2 represents the additive genetic, permanent environmental, and residual variance, respectively.

Genetic and phenotypic correlations were computed based on the estimated (co)variance components with genetic correlations being defined as the ratio of genetic covariances to the square root of the product of the corresponding variances and, similarly, phenotypic correlations being defined as the ratio of phenotypic covariances to the square root of the product of the corresponding variances. The required phenotypic covariance and variance are obtained by summing up the relevant covariance components.

Because the MTC program used EM-REML, no standard errors were directly available. At convergence, the approximate standard errors (SE) of all estimated parameters were determined using the algorithm by Meyer and Houle (2013), as implemented in the BLUPF90+ software (version 2.48), through the computation of the average information matrix.

2.4. Quantifying the ability of the 19 other traits to genetically predict LPNEB

The abilities of the other traits to predict genetically LPNEB were investigated using selection index theory (Van Vleck, 1993), by partitioning the genetic variance for LPNEB into contributions from the other 19 traits. This was done using the estimated genetic correlation matrix for dataset II representing standardized covariances. The partitioning process was conducted as follows for each trait or set of traits (i.e., LEDS and the other 18 traits):

1. Standardized genetic covariances (i.e., correlations) between LPNEB and the identified trait(s) were extracted and organized into a vector **c**;

2. Standardized genetic covariances among the identified traits were extracted and compiled into a matrix **G**;

3. A vector of regression coefficients **b** was obtained as $(\mathbf{c'G^{-1}})'$;
4. Explained variance was computed as $\mathbf{b'Gb = c'G^{-1}GG^{-1}c = c'G^{-1}c;}$
5. Finally, the genetic variance was partitioned into explained and unexplained components by comparing the explained variance to the total variance (assumed here to be equal to 1).

2.5. Causal effects of predicted EB on other 19 traits

The (co)variance components estimated from the multi-trait model can be transformed posterior into the (co)variance decomposition of a recursive model (Varona and González-Recio, 2023) to detect the causal effects. During this transformation process, the structural coefficients of LPNEB for the remaining 19 traits were computed. Furthermore, we assessed the causal effect of LPNEB on these 19 traits by analyzing the percentage change in genetic (or phenotypic) variance between the multi-trait model and the recursive model. Additional data preparation and processing were done using R (version 4.2.1, https://www.r-project.org/).

3. Results and discussion

3.1. Descriptive statistics

The descriptive statistics of LRNEB, LPNEB, LEDS, 15 biomarkers, and 3 production traits for the 2 datasets from first-parity cows are presented in Table 2-2. The predicted values for all traits used in this study were within the range of the measured data used previously for the initial calibration (i.e., model developments) of the MIR prediction models for traits. In dataset I, it needs to be noted that mean raw data for EB showed values of -3.11 Mcal/d for measured EB and -3.05 for predicted EB with SD of 6.26 and 4.35, respectively. In dataset II the mean was 0.01 for predicted EB with a SD of 3.70 highlighting the differences between the two datasets. As explained earlier, comparison of descriptive parameters of LRNEB, LPNEB, LEDS across traits and data sets should be considered with great caution. The averages of 4 predicted blood traits LB_BHBA, NEFA, LIGF-1, and GLU in dataset II were similar to those found in dataset I. The average predicted B_BHBA, NEFA, and GLU in dataset II were consistent with the measured data of these traits from other studies (Kostensalo et al., 2023; Mota et al., 2023).

Table 2-2. Mean and standard deviation (SD) values for the 21 traits[1] used in dataset I and dataset II.

		Dataset I						Dataset II	
		1st parity (n=243)		All parities (n=965)		Adjusted to 1st parity		1st parity (n=30,634)	
Traits	Unit	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LRNEB	n/a[2]	-0.03	0.28	0.09	0.31	-0.03	0.30	N/A[3]	N/A[3]
LPNEB	n/a[2]	-0.50	0.35	-0.30	0.35	-0.50	0.31	-0.09	0.29
LEDS	n/a[2]	-0.57	0.15	-0.59	0.15	-0.57	0.15	-0.53	0.16
In blood									
LB_BHBA	log (mmol/L)	-0.33	0.15	-0.29	0.15	-0.33	0.15	-0.24	0.14
NEFA	µekv/L	490.97	277.42	546.23	244.56	490.97	242.30	474.36	225.22
LIGF-1	log (mg/L)	2.19	0.17	2.01	0.22	2.19	0.17	2.18	0.14
GLU	mmol/L	3.83	0.25	3.65	0.25	3.83	0.22	3.72	0.21
In milk									
LM_BHBA	log (mmol/L)	2.21	0.11	2.18	0.14	2.21	0.14	2.20	0.13
CIT	mmol/L	8.80	1.30	9.06	1.69	8.80	1.68	8.88	1.45
LACE	log (mmol/L)	-1.15	0.22	-1.18	0.21	-1.15	0.21	-1.10	0.20
C10:0	g/dL of milk	0.09	0.02	0.10	0.02	0.09	0.02	0.09	0.02
C14:0	g/dL of milk	0.42	0.08	0.44	0.08	0.42	0.08	0.40	0.07
C16:0	g/dL of milk	1.22	0.22	1.22	0.24	1.22	0.24	1.10	0.20
C18:0	g/dL of milk	0.51	0.09	0.51	0.10	0.51	0.10	0.48	0.09
C18:1cis-9	g/dL of milk	0.89	0.22	0.87	0.24	0.89	0.23	0.91	0.23
SCFA	g/dL of milk	0.34	0.06	0.35	0.06	0.34	0.06	0.32	0.06
MCFA	g/dL of milk	1.93	0.34	1.96	0.36	1.93	0.36	1.77	0.30
LCFA	g/dL of milk	1.82	0.39	1.80	0.40	1.82	0.40	1.81	0.38
FP	% milk	4.00	0.54	4.00	0.61	4.00	0.60	3.79	0.54
PP	% milk	3.18	0.31	3.14	0.31	3.18	0.31	3.13	0.29
MY	kg/d	25.32	6.54	35.58	9.37	25.32	7.10	26.94	5.33

[1]LRNEB = logarithm measured probability negative energy balance; LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB = log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood IGF-1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

[2]n/a = not applicable because trait has no units.

[3]N/A = not applicable because data were not available.

The mean values of predicted 11 milk biomarkers and PP in both used datasets were similar, and the averages predicted 11 milk biomarkers values were in line with previous studies (Grelet et al., 2016, 2024; Soyeurt et al., 2006, 2011). The datasets used in the current study are completely independent from those cited in the referenced studies. The mean MY, after correction for parity, were similar between the two datasets in this study, which is also consistent with the findings in the larger dataset of Atashi et al. (2024) used representing the Walloon dairy cattle population. The value for FP in dataset I was higher (4.00%) compared to dataset II (3.79%). This fact and differences in EB described above, could indicate differences between datasets in energy intake and fiber in the diet.

3.2. Pearson correlation of measured EB, predicted EB with other 19 traits

The r between LRNEB (only present in dataset I) and LPNEB with the other 19 traits in dataset I and dataset II are shown in Figure 2-3 and Table A-1. Relevant *P*-values can be found in the Table A-2.

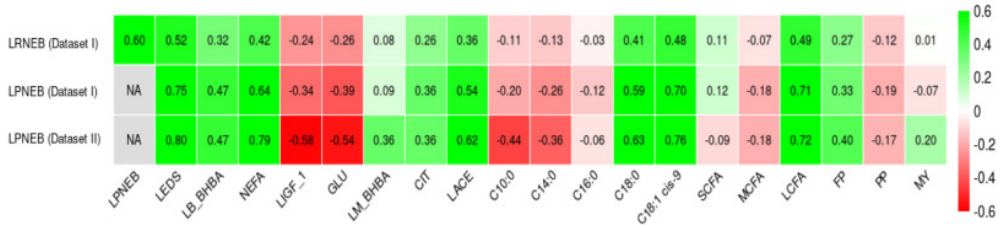


Figure 2-3. Comparison of Pearson correlations observed for LRNEB and LPNEB with the 19 other traits in dataset I (n = 965) and for LPNEB with the 19 other traits in dataset II (n = 30,634). LRNEB = logarithm measured probability NEB; LPNEB = logarithm probability NEB predicted by mid-infrared spectra; LEDS = logarithm probability EDS; LB_BHBA = log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; NEFA = blood nonesterified fatty acids predicted by mid-infrared spectra; LIGF-1 = log10 blood IGF-1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHBA = log10 milk β-hydroxybutyric acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium-chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield;
NA = not applicable because data not available.

The r coefficient computed between LRNEB and LPNEB in dataset I was 0.60. reflecting the R^2cv known from the calibration process (Table 2-1). This indicates a moderate but significant relationship between LRNEB and LPNEB. In the same dataset I, the r calculated between LRNEB and LEDS was 0.52, suggesting a comparatively slightly weaker association between LRNEB and LEDS. This finding aligns with expectations, as the predictive equation for EDS was not directly derived from EB, in contrast to prediction of EB. The absolute differences in r between LRNEB and the other traits with the corresponding values between LPNEB and the other traits are given in Figure 2-4. This low value of 0.13 suggests that despite moderate correlations between LRNEB and LPNEB these two traits showed a similar relationship with the other studied traits, indicating that despite working with MIR predicted traits some conclusions also valid for LRNEB can be drawn. This preliminary hypothesis was confirmed by comparing r for LRNEB (dataset I) and LPNEB (dataset II), which were similar to the values reported before and showed a mean absolute difference of 0.21.

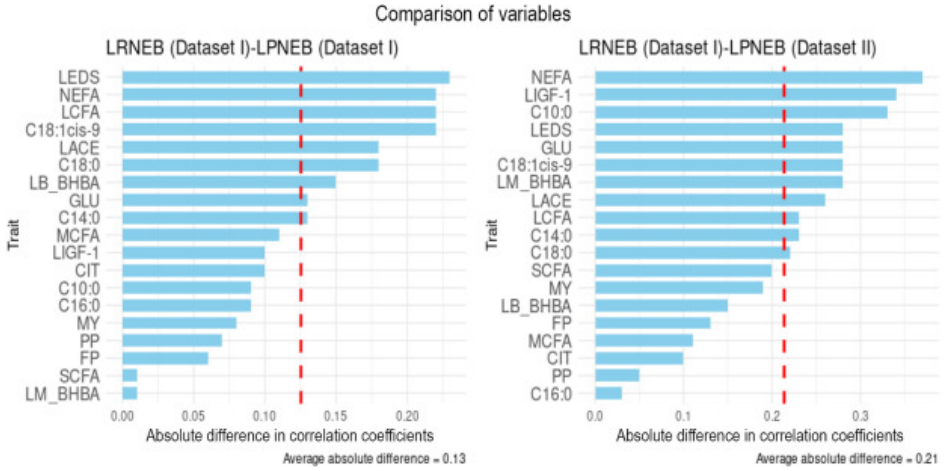


Figure 2-4. Absolute differences between Pearson correlations comparing LRNEB (dataset I) with LPNEB from datasets I and II for the other 19 traits. LEDS = logarithm probability EDS; LB_BHBA = log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; NEFA = blood nonesterified fatty acids predicted by mid-infrared spectra; LIGF-1 = log10 blood IGF-1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHBA = log10 milk beta-hydroxybutyric acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium-chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP =

milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

When comparing to literature, our results in dataset I for LRNEB with blood NEFA (0.42) were in line, but slightly lower, compared to those reported by Pires et al. (2022) for directly measured NEFA and EB. Our results indicated that fatty acids C18:1 cis-9 (0.48) and C18:0 (0.41) showed a strong r with LRNEB compared to the other milk biomarkers examined. These findings align with the results reported by Churakov et al. (2021), which also indicated that C18:1 cis-9 and C18:0 were correlated with NEB. The r calculated between LCFA and LRNEB was 0.49 which is in line with literature (e.g., Guinguina et al., 2021).

3.3. Genetic parameters

At convergence, relative off-diagonals were 0.42×10^{-3} for permanent environment and 0.11×10^{-3} for genetic (co)variances showing an excellent level of diagonalization. Figure 2-5 shows the h^2 (diagonal), genetic correlations (above diagonal), and phenotypic correlations (below diagonal) of the 20 traits examined (LPNEB, LEDS, 15 biomarkers, and 3 production traits) estimated by a 20-trait repeatability model. All the studied traits (n = 20) had a moderate h^2 (ranged from 0.16 to 0.29) and a SE ranging from 0.01 to 0.02, except for CIT which had a higher h^2 (0.38) with a SE of 0.02 (Table A-3). All h^2 were significantly different from zero (Table A-4) The h^2 of LPNEB was similar to those reported in previous studies with directly measured EB (McParland et al., 2015; Becker et al., 2021). Becker et al., (2021) showed that the h^2 of measured EB was not stable during the lactation, ranging from 0.15 to 0.31. The h^2 of 4 examined predicted blood biomarkers (LB_BHBA, NEFA, LIGF-1, and GLU) agreed with those reported by van der Drift et al., (2012), Benedet et al. (2020), and Hayhurst et al. (2009). However, Hayhurst et al. (2009) showed that the h^2 of IGF-1 varied in different populations, ranging from 0.21 to 0.66. The h^2 of LM_BHBA, CIT, and LACE were similar to those reported with a larger dataset (Ranaraja et al., 2018; Chen et al., 2023). The h^2 of the other 8 milk biomarkers and 3 production traits were within the range estimated by random regression models (Bastin et al., 2011; Paiva et al., 2022).

The absolute genetic correlation between LPNEB and 9 out of the remaining 19 traits (LEDS, LB_BHBA, NEFA, LIGF-1, GLU, LACE, C18:0, C18:1 cis-9, and LCFA) were higher than 0.50, and LPNEB has the highest genetic correlation with NEFA (0.87). In early lactation, blood NEFA has been considered by previous studies to be the best single biomarker of NEB (Oikonomou et al., 2008; Andjelić et al., 2022).

As it was mentioned previously (Franceschini et al., 2022), EDS included EB explaining the high correlation of 0.85 with LPNEB. Moreover, EDS included also C18:1 cis-9, so it showed the second-highest genetic correlation (0.81). The third-highest genetic correlation was found between LPNEB and C18:1 cis-9 (0.76).

Recently, Churakov et al. (2021) and Guinguina et al. (2021) reported that the Oleic acid is a good biomarker for identifying NEB state in dairy cows. We also found that the genetic correlation values between LPNEB, and PP and MY were -0.47 and 0.45, respectively. According to Buttchereit et al. (2011), the genetic correlations between direct EB and energy-corrected milk yield in early lactation (15–45 DIM) ranged from –0.57 to –0.39, whereas those between direct EB and PP ranged from 0.30 to 0.53.

The genetic correlation values among the other 18 traits are also within a reasonable range. For example, the genetic correlation estimated among NEFA and LM_BHBA, LACE, C18:1 cis-9 were similar to that reported by Mehtiö et al. (2020). The 8 examined fatty acids (FA) had high genetic correlation with FP (> 0.50), which were similar to results reported by Bastin et al. (2011).

The phenotypic correlations estimated between each pair of the 20 traits were generally similar to the corresponding genetic correlations, with the phenotypic correlations being slightly lower than the genetic correlations. The average absolute difference of genetic correlations and phenotypic correlations for LPNEB with the other 19 traits was 0.06 with a minimum of 0.00 for C14:0, LCFA and a maximum of 0.17 for MY. The average absolute difference of genetic correlations and phenotypic correlations for LEDS with the other 19 traits was 0.06 with a minimum of 0.00 for LACE, C18:1cis-9 and a maximum of 0.21 for MY.

The absolute phenotypic correlations between LPNEB and 8 traits (LEDS, LB_BHBA, NEFA, LIGF-1, LACE, C18:0, C18:1 cis-9, and LCFA) were higher than 0.50, including the phenotypic correlation of LEDS was the highest (0.81). The phenotypic correlations between LPNEB and NEFA, C18:1 cis-9 were 0.78 and 0.70, respectively. The high phenotypic correlations of these 3 traits were consistent with the genetic correlations mentioned above. Severe NEB state in cattle expressed by high LPNEB will increase 7 of the 8 traits (LEDS, LB_BHBA, NEFA, LACE, C18:0, C18:1 cis-9, LCFA) and decrease 1 biomarker (LIGF-1).

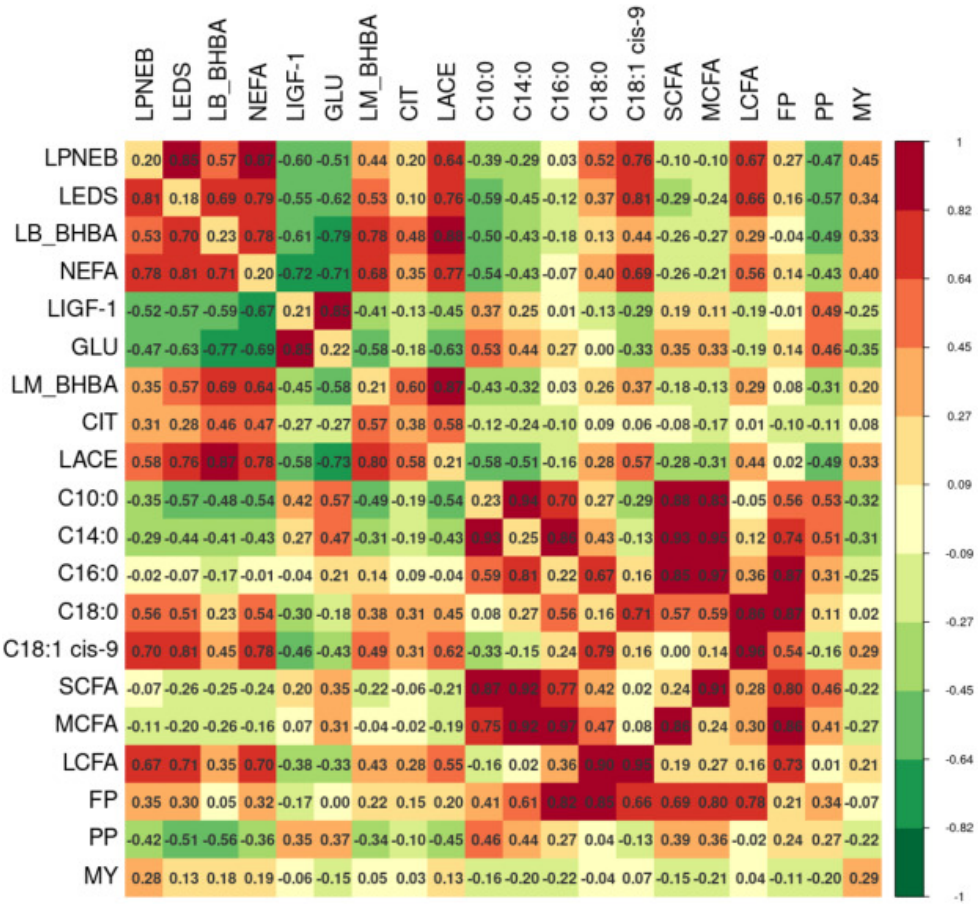


Figure 2-5. The heritability (diagonal) of 20 traits and their genetic (above diagonal), and phenotypic (below diagonal) correlations. The ranges of SE for h^2, genetic correlation, and phenotypic correlation are from 0.01 to 0.02, 0.01 to 0.07, and below 0.01 to 0.06, respectively. LPNEB = logarithm probability NEB predicted by mid-infrared spectra; LEDS = logarithm probability EDS; LB_BHBA = log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; NEFA = blood nonesterified fatty acids predicted by mid-infrared spectra; LIGF-1 = log10 blood IGF-1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHBA = log10 milk β-hydroxybutyric acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium-chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

3.4. Quantifying the ability of the 19 other traits to genetically predict LPNEB

The quantification of the ability of the other 19 traits to predict genetically LPNEB, expressed as the fraction of the genetic variance of LPNEB that could be predicted by other traits, was computed to be 65% for LEDS alone, 62% for NEFA alone, 82% for all 15 biomarkers alone, 85% for all other 18 traits excluding LEDS and 89% combining LEDS and the other traits. These results seem to indicate that LEDS contains additional information useful to predict LPNEB not contained in the other 18 traits. This result could be due to the specific approach (i.e., cluster analysis) used by Franceschini et al. (2022) leading to the definition of a specific NEB related cluster (i.e., his Cluster 4), which is the bases of EDS, respectively LEDS, as used in this study.

3.5. Causal effects of predicted EB on other 19 traits

The structural coefficients and causal effects (phenotypic and genetic variances changed percentage) of the LPNEB on the other 19 traits are shown in Table 2-3. Structural coefficients represent the amount of change in trait i with one unit change of LPNEB expressed as decrease in trait *i* (i.e., which is equivalent to expressing trait *i* but keeping LPNEB constant). Positive values for structural coefficients for NEFA, CIT, and MY mean that higher values of LPNEB lead to higher values for these traits. However, as it is well known, the NEB state of cattle is caused by insufficient energy intake to produce MY, therefore the causality of MY and NEB state might be reversed (i.e., changes in MY lead to higher NEB) even if the fact that animals are in a state of NEB leads to less MY. The NEFA can measure fat mobilization in cattle, which is an important indicator of NEB state in cattle (Zachut et al., 2020). Citrate is implicated in this process and is an inhibitor of de novo FA synthesis in dairy cows (Garnsworthy et al., 2006). Citrate is also considered a potential early biomarker for assessing NEB state in dairy cows (Bjerre-Harpøth et al., 2012; Xu et al., 2020).

Table 2-3 also shows phenotypic and genetic variance changes due to fitting the recursiveness to LPNEB with positive values indicating a decrease and negative values an increase. In total 8 traits (LEDS, LB_BHBA, NEFA, LIGF-1, LACE, C18:1 cis-9, LCFA and MY) were selected based on the results of phenotypic and genetic variance change (both absolute values higher than 30%). The top-3 of the 8 traits were NEFA, LEDS, C18:1 cis-9, which are the same as the genetic correlations. The phenotypic (positive) and genetic (negative) variance change of MY were in different directions, which again supports the idea that MY and LPNEB had an inversed causality.

Based on our results (r, genetic and phenotypic correlations, explained variances, structural coefficients and caused effects) we can suggest 8 out of the 19 examined

traits (LEDS, LB_BHBA, NEFA, LIGF-1, CIT, LACE, C18:1 cis-9, LCFA) as being highly related to LPNEB. Our results indicated that the most relevant biomarker was NEFA, which is already well-recognized in literature (e.g., Macrae et al., 2009). However, NEFA is also one of MIR-based biomarkers based on the poorest calibrations (Table 2-1), therefore we want to express here some caution. The observed recursive relationships between 8 traits with LPNEB are summarized in Figure 2-6, indicating that high LPNEB will impact the other 7 individual biomarkers in the direction indicated by the arrows. Moreover, based on its definition, LEDS represents 4 (LB_BHBA, CIT, LACE, C18:1 cis-9) of the major single biomarkers considered impacted by NEB state.

Table 2-3. Structural coefficients and causal effects (phenotypic and genetic variances changed percentage) of log10 predicted negative energy balance (LPNEB) on the 19 traits[1]

Traits	Structural coefficients	Phenotypic variance change (%)	Genetic Variance change (%)
LEDS	-0.44	73.66	71.51
In blood			
LB_BHBA	-0.36	38.89	30.88
NEFA	-6.86	81.06	75.43
LIGF-1	0.37	46.25	33.22
GLU	0.60	35.32	17.29
In milk			
LM_BHBA	-0.23	25.51	18.30
CIT	-1.57	7.75	4.18
LACE	-0.50	47.19	39.79
C10:0	0.05	19.82	12.55
C14:0	0.15	15.58	5.32
C16:0	0.17	-0.20	-5.98
C18:0	-0.07	14.98	18.97
C18:1 cis-9	-0.42	62.37	54.83
SCFA	0.07	3.54	-3.85
MCFA	0.41	4.45	-3.94
LCFA	-0.53	42.26	40.14
FP	-0.03	0.41	0.73
PP	0.70	36.90	22.88
MY	-2.38	40.59	-31.53

[1]LEDS = logarithm probability energy deficit score; LB_BHBA= log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra;

LM_BHBA = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

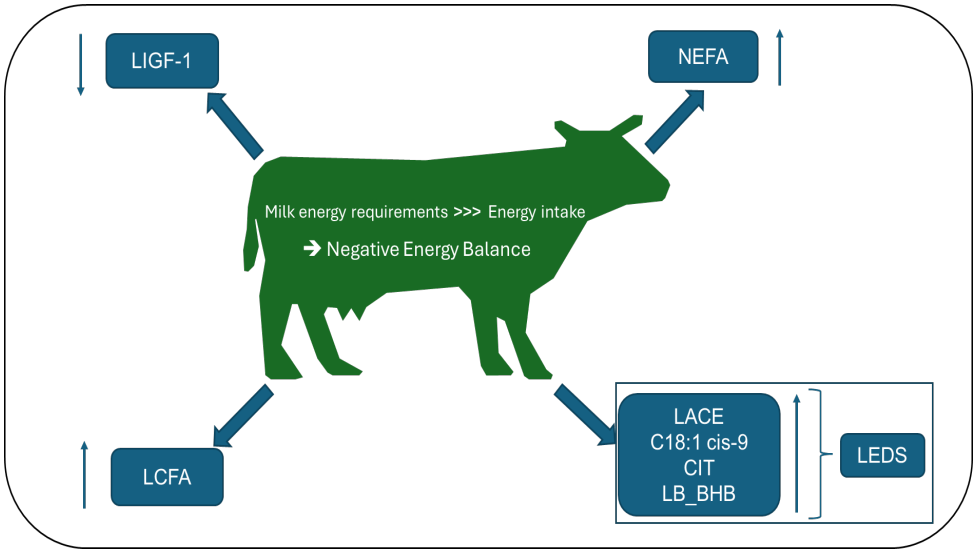


Figure 2-6. Reaction of 7 important biomarkers as reported in this study when cows had negative energy balance in early lactation (LIGF-1 = log10 blood IGF-1 predicted by mid-infrared spectra; NEFA = blood nonesterified fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LB_BHBA = log10 blood β-hydroxybutyric acid predicted by mid-infrared spectra; LEDS = logarithm probability EDS). The arrows represent the direction of change in trait values when dairy cows are in NEB. The upward arrow indicates an increase in the trait value, and the downward arrow indicates a decrease in the trait value.

4. Conclusions

This study explored the relationships among LRNEB, LPNEB, LEDS, 15 biomarkers, and 3 production traits. The LPNEB showed a strong genetic correlation with NEFA, confirming its relevance as a biomarker, despite its lower prediction accuracy. The genetic correlation patterns between LRNEB and LPNEB with other

traits were similar, suggesting LPNEB can be cautiously used in place of LRNEB. The novel trait LEDS exhibited the strongest causal association with NEB and demonstrated a high genetic ability to predict LPNEB (65%), surpassing NEFA (62%). A recursive model identified strong causal links between predicted NEB and 8 traits, also strengthening the case for using LEDS as an indicator of NEB. These findings enhance our understanding of the genetic, phenotypic, and causal relationships of NEB with the other 19 studied traits. Further research is needed to explore the genetic architecture of these traits and their relationships.

5. Acknowledgments

The China Scholarship Council (Beijing, China) is acknowledged for funding the PhD project of Hongqing Hu. The providing of a dataset collected by the GplusE Project (https://gpluse.eaap.org/) on experimental farms is acknowledged. 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 provided also support through the PDR projects "HTwoTHI" (grant number T.W005.23) and "DEEPSELECT" (grant number T.0095.19). Supplemental material for this article is available at https://data.mendeley.com/datasets/gh9pwf8fph/2. Author contributions are as follows: Hongqing Hu, funding acquisition, formal analysis, and writing (original draft); Sébastien Franceschini, data curation and writing (review and editing); Pauline Lemal, validation and writing (review and editing); Clément Grelet, data curation and writing (review and editing); Yansen Chen, data curation and writing (review and editing); Hadi Atashi, validation and writing (review and editing); Katrien Wijnrocx, validation and writing (review and editing); Hélène Soyeurt, validation and writing (review and editing); Nicolas Gengler, conceptualization, funding acquisition, methodology, project administration, software, supervision, validation, and writing (review and editing). 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.

6. References

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.

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.

Baumgard, L.H., L.J. Odens, J.K. Kay, R.P. Rhoads, M.J. VanBaale, and R.J. Collier. 2006. Does negative energy balance (NEBAL) limit milk synthesis in early lactation. In Proc. Southwest Nutr. Conf (pp. 181-187). https://rapp.ualberta.ca/wpcontent/uploads/sites/57/wcds_archive/Archive/2007/Manuscripts/Lance.pdf.

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.

Berry, D.P., B. Horan, M. O'Donovan, F. Buckley, E. Kennedy, M. McEvoy, and P. Dillon. 2007. Genetics of grass dry matter intake, energy balance, and digestibility in grazing Irish dairy cows. J. Dairy Sci. 90:4835–4845. https://doi.org/10.3168/jds.2007-0116.

Bjerre-Harpøth, N.C. Friggens, V.M. Thorup, T. Larsen, B.M. Damgaard, K.L. Ingvartsen, and K.M. Moyes. 2012. Metabolic and production profiles of dairy cows in response to decreased nutrient density to increase physiological imbalance at different stages of lactation. J. Dairy Sci. 95:2362-2380. https://doi.org/10.3168/jds.2011-4419.

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., S. Vanderick, R.R. Mota, C. Grelet, GplusE Consortium, and N. Gengler. 2021. Estimation of genetic parameters for predicted nitrogen use efficiency and losses in early lactation of Holstein cows. J. Dairy Sci. 104:4413–4423. https://doi.org/10.3168/jds.2020-18849.

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.

Collard, B.L., P.J. Boettcher, J.C.M. Dekkers, D. Petitclerc, and L.R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early

lactation. J. Dairy Sci. 83:2683–2690. https://doi.org/10.3168/jds.S0022-0302(00)75162-9.

Franceschini, S., C. Grelet, J. Leblois, N. Gengler, H. Soyeurt, and GplusE Consortium. 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.

Garnsworthy, P.C., L.L. Masson, A.L. Lock, and T.T. Mottram. 2006. Variation of milk citrate with stage of lactation and de novo fatty acid synthesis in dairy cows. J. Dairy Sci. 89:1604–1612. https://doi.org/10.3168/jds.S0022-0302(06)72227-5.

Rustin, M., S. Janssens, N. Buys, and N. Gengler. 2009. Multi-trait animal model estimation of genetic parameters for linear type and gait traits in the Belgian warmblood horse. J. Anim. Breed. Genet. 126:378-386. https://doi.org/10.1111/j.1439-0388.2008.00798.

Grelet, C., A. Vanlierde, F. Dehareng, E. Froidmont, and GplusE 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, Wageningen Academic Publishers. p. 403. 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, GplusE 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.

Grelet, C., C. Bastin, M. Gelé, J.B. Davière, M. Johan, A. Werner, R. Reding, J.A. Fernandez Pierna, F.G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825. https://doi.org/10.3168/jds.2015-10477.

Grelet, C., E. Froidmont, L. Foldager, M. Salavati, M. Hostens, C.P. Ferris, K.L. Ingvartsen, M.A. Crowe, M.T. Sorensen, J.A. Fernandez Pierna, A. Vanlierde, N. Gengler, GplusE Consortium, and F. Dehareng,. 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. J. Dairy Sci. 103:4435–4445. https://doi.org/10.3168/jds.2019-17910.

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., P. Dardenne, H. Soyeurt, J.A. Fernandez, A. Vanlierde, F. Stevens, N. Gengler, and F. Dehareng. 2021. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. Methods 186:97–111. https://doi.org/10.1016/j.ymeth.2020.07.012.

Grelet, C., T. Larsen, M.A. Crowe, D.C. Wathes, C.P. Ferris, K.L. Ingvartsen, C. Marchitelli, F. Becker, A. Vanlierde, J. Leblois, U. Schuler, F.J. Auer, A. Köck, L. Dale, J. Sölkner, O. Christophe, J. Hummel, A. Mensching, J.A.F. Pierna, H. Soyeurt, M. Calmels, R. Reding, M. Gelé, Y. Chen, N. Gengler, and F. Dehareng. 2024. Prediction of key milk biomarkers in dairy cows through milk MIR spectra and international collaborations. J. Dairy Sci. 107:1668-1684. https://doi.org/10.3168/jds.2023-23843.

Guinguina, A., T. Yan, E. Trevisi, and P. Huhtanen. 2021. The use of an upgraded GreenFeed system and milk fatty acids to estimate energy balance in early-lactation cows. J. Dairy Sci. 104:6701–6714. https://doi.org/10.3168/jds.2020-19591.

Hayhurst, C., A.P.F. Flint, P. Løvendahl, J.A. Woolliams, and M.D. Royal. 2009. Genetic variation of metabolite and hormone concentration in UK Holstein-Friesian calves and the genetic relationship with economically important traits. J. Dairy Sci. 92:4001–4007. https://doi.org/10.3168/jds.2008-1130.

Heirbaut, S., X.P. Jing, B. Stefańska, E. Pruszyńska-Oszmałek, L. Buysse, P. Lutakome, M.Q. Zhang, M. Thys, L. Vandaele, and V. Fievez. 2023. Diagnostic milk biomarkers for predicting the metabolic health status of dairy cattle during early lactation. J. Dairy Sci. 106:690–702. https://doi.org/10.3168/jds.2022-22217.

Ho, P.N., L.C. Marett, W.J. Wales, M. Axford, E.M. Oakes, and J.E. Pryce. 2020. 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.

Kostensalo, J., M. Lidauer, B. Aernouts, P. Mäntysaari, T. Kokkonen, P. Lidauer, and T. Mehtiö. 2023. Short communication: Predicting blood plasma non-esterified fatty acid and beta-hydroxybutyrate concentrations from cow milk—addressing systematic issues in modelling. Animal 17:100912. https://doi.org/10.1016/j.animal.2023.100912.

ICAR. 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.

Liinamo, A.E., P. Mäntysaari, and E.A. Mäntysaari. 2012. Short communication: Genetic parameters for feed intake, production, and extent of negative energy balance in Nordic Red dairy cattle. J. Dairy Sci. 95:6788–6794. https://doi.org/10.3168/jds.2012-5342.

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.

McParland, S., and D.P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J. Dairy Sci. 99:4056–4070. https://doi.org/10.3168/jds.2015-10051.

McParland, S., E. Kennedy, E. Lewis, S.G. Moore, B. McCarthy, M. O'Donovan, and D.P. Berry. 2015. Genetic parameters of dairy cow energy intake and body energy status predicted using mid-infrared spectrometry of milk. J. Dairy Sci. 98:1310–1320. https://doi.org/10.3168/jds.2014-8892.

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.

Meyer, K., and D. Houle. 2013. Sampling based approximation of confidence intervals for functions of genetic covariance matrices. Proc. Assoc. Advmt. Anim. Breed. Genet. 20: 523-526. https://hdl.handle.net/1959.11/14324.

Mota, L.F.M., D. Giannuzzi, S. Pegolo, E. Trevisi, P. Ajmone-Marsan, and A. Cecchinato. 2023. Integrating on-farm and genomic information improves the predictive ability of milk infrared prediction of blood indicators of metabolic disorders in dairy cows. Genet. Sel. Evol. 55:1–18. https://doi.org/10.1186/s12711-023-00795-1.

Misztal, I., Weigel, K., and T.J. Lawlor. 1995. Approximation of estimates of (co) variance components with multiple-trait restricted maximum likelihood by multiple diagonalization for more than one random effect. J. Dairy Sci. 78: 1862-1872. https://doi.org/10.3168/jds.S0022-0302(95)76811-4.

Oikonomou, G., G. Arsenos, G.E. Valergakis, A. Tsiaras, D. Zygoyiannis, and G. Banos. 2008. Genetic relationship of body energy and blood metabolites with reproduction in Holstein cows. J. Dairy Sci. 91:4323–4332. https://doi.org/10.3168/jds.2008-1018.

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.

Pires, J.A.A., T. Larsen, and C. Leroux. 2022. Milk metabolites and fatty acids as noninvasive biomarkers of metabolic status and energy balance in early-lactation cows. J. Dairy Sci. 105:201–220. https://doi.org/10.3168/jds.2021-20465.

Ranaraja, U., K.H. Cho, M.N. Park, S.D. Kim, S.H. Lee, and C.H. Do. 2018. Genetic parameter estimation for milk β-hydroxybutyrate and acetone in early lactation and its association with fat to protein ratio and energy balance in Korean Holstein cattle. Asian-Australasian J. Anim. Sci. 31:798–803. https://doi.org/10.5713/ajas.17.0443.

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.

Stádník L., F. Louda, and A. Jezkova. 2002. The effect of selected factors at insemination on reproduction of Holstein cows. Czech J. Anim. Sci. 47: 169-175.

Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating Fatty Acid Content in Cow Milk Using Mid-Infrared Spectrometry. J. Dairy Sci. 89:3690–3695. https://doi.org/10.3168/jds.s0022-0302(06)72409-2.

Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D.P. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. J. Dairy Sci. 94:1657–1667. https://doi.org/10.3168/jds.2010-3408.

van der Drift, S.G.A., K.J.E. van Hulzen, T.G. Teweldemedhn, R. Jorritsma, M. Nielen, and H.C.M. Heuven. 2012. Genetic and nongenetic variation in plasma and milk β-hydroxybutyrate and milk acetone concentrations of early-lactation dairy cows. J. Dairy Sci. 95:6781–6787. https://doi.org/10.3168/jds.2012-5640.

Van Vleck, L. D. 1993. Selection index and introduction to mixed model methods for genetic improvement of animals: The green book. CRC Press, Boca Raton, FL.

Varona, L., and O. González-Recio. 2023. Invited review: Recursive models in animal breeding: Interpretation, limitations, and extensions. J. Dairy Sci. 106:2198–2212. https://doi.org/10.3168/jds.2022-22578.

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., L. L. M.Thornton, R. R. Neitzel, and N. Gengler. 2006. Genetic parameters and evaluation of rear legs (rear view) for Brown Swiss and Guernseys. J. Dairy Sci. 89: 4895-4900. https://doi.org/10.3168/jds.S0022-0302(06)72538-3.

Xu, W., J. Vervoort, E. Saccenti, B. Kemp, R.J. van Hoeij, and A.T.M. van Knegsel. 2020. Relationship between energy balance and metabolic profiles in plasma and milk of dairy cows in early lactation. J. Dairy Sci. 103:4795–4805. https://doi.org/10.3168/jds.2019-17777.

Zachut, M., M. Šperanda, A.M. De Almeida, G. Gabai, A. Mobasheri, 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/S00220299200000.

3

Chapter III Comparing genetic architecture of MIR-predicted energy balance, a novel energy deficiency score and several biomarkers

Adapted from: Hu, H., S. Franceschini, P. Lemal, H. Atashi, C. Grelet, Y. Chen, K. Wijnrocx, H. Soyeurt, and N. Gengler. 2025. Comparing genetic architecture of MIR-predicted energy balance, a novel energy deficiency score and several biomarkers. J. Dairy Sci. https://doi.org/10.3168/jds.2025-26876. Accepted.

Foreword

Dairy cows often experience negative energy balance (NEB) during early lactation, which can negatively impact their health and milk production. This study compared the genetic architectures of logit-transformed predicted NEB (LPNEB), logit-transformed novel energy deficiency score (LEDS), and analyzed their genetic relationships with 15 biomarkers and 3 production traits. Blood non-esterified fatty acids (NEFA) showed strong genomic correlations with both LPNEB and LEDS, suggesting direct roles in energy metabolism. Key genomic regions on BTA25 and BTA29 were linked to fat mobilization and ketone synthesis. These findings improve our understanding of NEB genetics and support genetic selection for enhanced metabolic performance.

Abstract

Negative energy balance (NEB) during early lactation is a critical physiological challenge in high-producing dairy cows, impacting 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. 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 non-esterified 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

1. 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 the 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 heritability 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 one or two biomarkers, such as non-esterified fatty acids (NEFA), beta-hydroxybutyrate acid (BHBA), 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, non-invasive, and scalable. Recent advances in mid-infrared (MIR) spectrometry have provided a promising, large-scale, and non-invasive 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, BHBA, and insulin-like growth factor 1 (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), which was derived from 27 MIR-predicted metabolic features, such as milk BHBA and oleic acid (C18:1cis-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 LEDS. In the study by Hu et al. (2025), LEDS showed strong genetic correlations with logit-transformed predicted NEB (LPNEB, 0.85), blood BHBA (0.69), and milk C18:1cis-9 (0.81), further supporting its potential

as a biologically relevant and genetically informative trait related to energy metabolism during early lactation. Moreover, the heritability 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 BHBA, 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.

2. Materials and methods

2.1. 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 were described by Hu et al. (2025) in their Table 1. The LPNEB was derived from the predicted energy balance (PEB) as follows:

$$\text{LPNEB} = \log_{10}\frac{PNEB}{1 - PNEB} \quad (3.1)$$

where $\text{PNEB} = 1 - \frac{\text{PEB} - \text{PEB}_{\text{minimum}}}{\text{PEB}_{\text{maximun}} - \text{PEB}_{\text{minimum}}}$, 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 were 0.43 (Grelet et al., 2017).

The LEDS was derived from the EDS as follows:

$$\text{LEDS} = \log_{10} \frac{EDS}{1 - EDS} \quad (3.2)$$

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 log10-transformed blood BHBA (LB_BHBA), blood NEFA, log10-transformed blood insulin-like growth factor 1 (LIGF-1), blood glucose (GLU), log10-transformed milk beta-hydroxybutyrate (LM_BHBA), milk citrate (CIT), log10-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 fatty acids (SCFA), milk medium chain fatty acids (MCFA), milk long-chain fatty acids (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 above mentioned. The MY, FP, and PP were limited from 3 to 99 kg/day, 1 to 9%, and 1 to 7%, respectively (ICAR, 2022).

The Global H measure was calculated based on standardized Mahalanobis distances between each standardized MIR spectrum in this study and the MIR spectra in the calibration dataset of the different traits (Whitfield et al., 1987). We only kept MIR records with Global H ≤ 3 for all MIR predictions. 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_BHBA).

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 Figure A-1. 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, USA). SNPs common across the four chips were retained for analysis. Non-mapped 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 (MAF) 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 Bos taurus autosome (BTA) were retained for genomic analysis.

2.2. Variance component estimation

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

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Mm} + \mathbf{Cc} + \mathbf{Dd} + \mathbf{Qq} + \mathbf{Wp} + \mathbf{Za} + \mathbf{e} \quad (3.3)$$

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 to 24 months, 25, 26, 27, 28, 29, 30, 31 to 32 months, 33 and 34 months, and 35 to 37 months); **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 to 6, DIM 7 to 8, DIM 9 to 10, and DIM 11 to 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; **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 following (Aguilar et al., 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad (3.4)$$

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:

$$\mathbf{G} = \mathbf{G}^* \times 0.95 + \mathbf{A}_{22} \times 0.05 \quad (3.5)$$

where **G*** represents the genomic relationship matrix as defined by the first method of VanRaden (2008).

Computations are 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 heritability (h^2) and repeatability (r) are as follows.

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2} \quad (3.6)$$

$$r = \frac{\sigma_a^2 + \sigma_p^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2} \quad (3.7)$$

where σ_a^2, σ_p^2 and σ_e^2 represents the additive genetic, permanent environmental, and residual variance, respectively.

2.3. SNP effect estimation

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

$$\widehat{\boldsymbol{\beta}} = \mathbf{D}\mathbf{Z}_\mathbf{g}'\left[\mathbf{Z}_\mathbf{g}\mathbf{D}\mathbf{Z}_\mathbf{g}'\right]^{-1}\widehat{\mathbf{u}} \quad (3.8)$$

where **D** = **I**, meaning all SNPs were given equal weight; **Zg** was an incidence matrix of genotyped for each SNP; $\widehat{\mathbf{u}}$ was a vector of GEBV for each trait.

2.4. Estimation of SNP-based genomic correlations

This analysis used the elements of the vector $\widehat{\boldsymbol{\beta}}$ 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:

$$Cov_{ij} = \sum_{k=1}^{n} 2p_k(1-p_k)\beta_{ik}\beta_{jk} \quad (3.9)$$

where $\boldsymbol{p}_\mathbf{k}$ is the allele frequency of the k^{th} SNP, $\boldsymbol{\beta}_{\mathbf{ik}}$ and $\boldsymbol{\beta}_{\mathbf{jk}}$ 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:

$$Var_i = \sum_{k=1}^{n} 2\boldsymbol{p_k}(1-\boldsymbol{p}_k)\boldsymbol{\beta}_{ik}^2 \quad (3.10)$$

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

$$r_{ij} = \frac{Cov_{ij}}{\sqrt{Var_i \cdot Var_j}} = \frac{\sum 2p(1-p)\cdot\beta_{ij}}{\sqrt{\sum 2p(1-p)\cdot\beta_i^2 \cdot \sum 2p(1-p)\cdot\beta_j^2}} \quad (3.11)$$

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:

$$SE(r) \approx \sqrt{\frac{(1-r^2)^2}{n-2}} \quad (3.12)$$

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 correlation coefficient r, the z-score was calculated as:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \quad (3.13)$$

The statistical significance of the difference between the chromosome-specific correlation (r_{BTA}) and the genome-wide correlation (r_{All}) was assessed using the following test statistic:

$$z_{diff} = \frac{z_{BTA} - z_{All}}{\sqrt{\frac{1}{n_{BTA}-3} + \frac{1}{n_{All}-3}}} \quad (3.14)$$

where n_{BTA} and n_{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:

$$MeanAbsCorr_{BTA} = \frac{\sum_{i=1}^{19} |r_{BTA,i}|}{19} \quad (3.15)$$

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

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

2.5. Estimation of the independent contributions of 8 traits to LPNEB and LEDS

A total of 8 traits LB_BHBA, NEFA, LIGF-1, LM_BHBA, 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 and biological interpretability.

To quantify the independent contributions of 8 traits to LPNEB and LEDS, we used a selection index-based approach using the SNP-based genomic correlations. This analysis was performed at both the whole-genome level (All_BTA) and across each of the 29 autosomes (BTA1 to 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:

$$\mathbf{w} = \mathbf{r}_{\text{target}}\mathbf{C}^{-1} \quad (3.16)$$

where $\mathbf{r}_{\text{target}}$ 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:

$$\text{Total Contribution} = \mathbf{wCw}^{\mathbf{T}} \quad (3.17)$$

Where $\mathbf{w} = \mathbf{r}_{\text{target}} \cdot \mathbf{C}^{-1}$ represents the vector of independent standardized contributions of the traits to 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. To examine 10 traits specific contributions on each chromosome, this procedure was applied separately to the SNP effects from each chromosome. 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:

$$\text{Relative Contribution}_i = \frac{w_i}{\sum_{j=1}^{n}|w_j|} \quad (3.18)$$

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).

3. Results and discussion

3.1. Genetic parameters

The h^2 and r estimates of the 20 traits are presented in Figure 3-1. Estimates obtained using the single-trait repeatability model were highly comparable to those derived from the multi-trait 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 to 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_BHBA, 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_BHBA, 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 fatty acid 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 a 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.

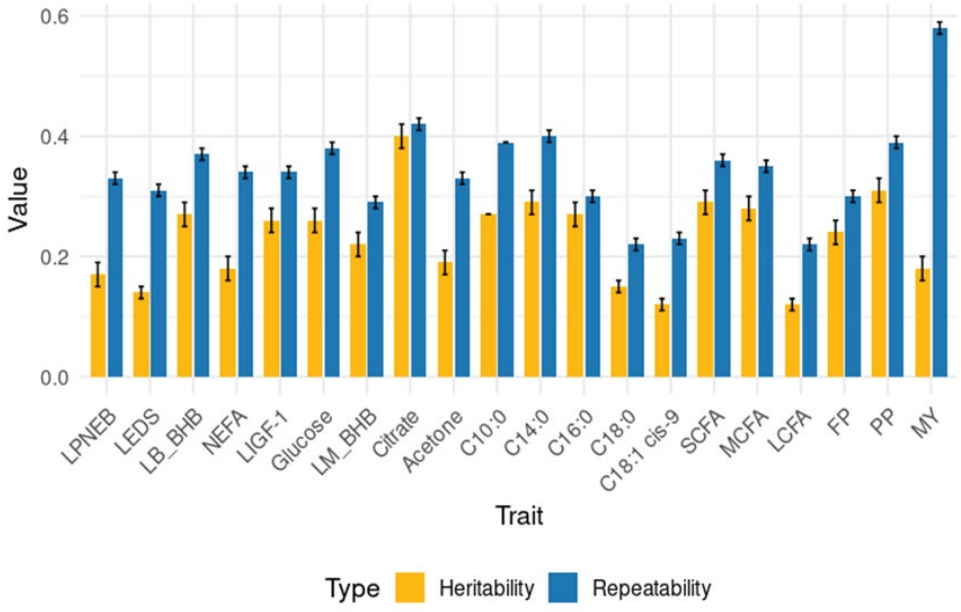


Figure 3-1. Heritability and repeatability of 20 traits estimated based on 20 univariate repeatability models. Error bars represent standard errors (0–0.02 for heritability; 0–0.01 for repeatability). LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHBA= log_{10}-transformed blood beta-hydroxybutyrate predicted by MIR spectra; NEFA= blood non-esterified fatty acids predicted by MIR spectra; LIGF-1= log_{10}-transformed blood insulin-like growth factor 1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHBA = log_{10}-transformed milk beta-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.

3.2. 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 3-2). 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 two modeling approaches. The estimates from both models ranged from -0.8 and -0.79 (GLU and LB_BHBA) 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 multi-trait model.

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_BHBA, NEFA, LIGF-1, LACE, C18:1 cis-9, and LCFA. 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; Mansour et al., 2022; Mehtiö et al., 2020). These strong correlations suggest that both LPNEB and LEDS effectively capture the underlying genetic architecture of this key metabolic process. LPNEB and LEDS showed strong genomic correlations with two major ketone bodies, LB_BHBA and LACE, which 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 BHBA and acetone (Guliński, 2021; Lisuzzo et al., 2022; Martens, 2023). C18:1 cis-9, a monounsaturated

long-chain fatty acid, showed strong positive correlations with both LPNEB and LEDS. Increased fat mobilization during NEB results in higher levels of long-chain fatty acids, including C18:1 cis-9, due to enhanced hepatic β-oxidation (Bastin et al., 2012; Churakov et al., 2021), which supports our findings.

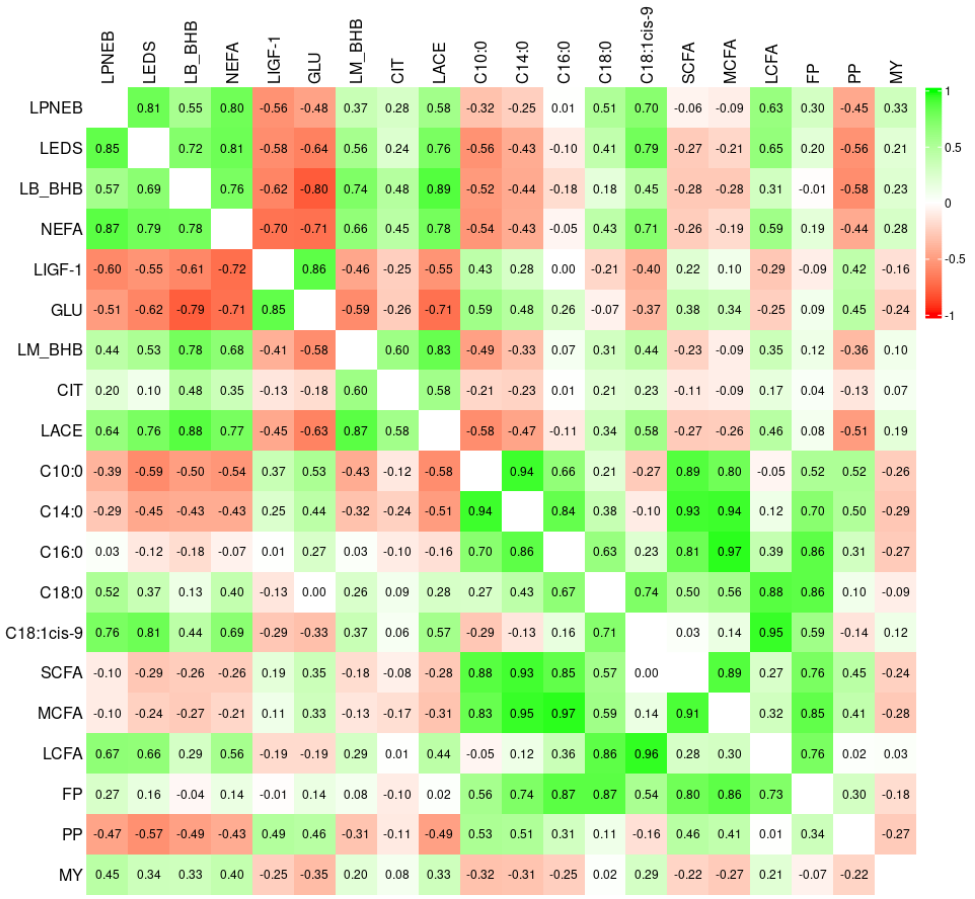


Figure 3-2. Genomic correlations among 20 traits. Values in the upper diagonal were calculated based on SNP effects from 20 univariate repeatability models, while values in the lower diagonal were derived from 20-traits repeatability model. The SE of the correlation matrix ranged from 0.0003 to 0.0013. The 20 traits include: LPNEB = logarithm probability negative energy balance predicted by mid-infrared (MIR) spectra; LEDS = logarithm probability energy deficiency score; LB_BHBA= log_{10}-transformed blood beta-hydroxybutyrate predicted by MIR spectra; NEFA= blood non-esterified fatty acids predicted by MIR spectra; LIGF-1= log_{10}-transformed blood insulin-like growth factor 1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHBA = log_{10}-transformed milk beta-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.

3.3. 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 3-3. The significance tests for the genomic correlations shown in Figure 3-3 are presented in Table A-5. At the All_BTA, LPNEB showed strong positive genetic correlations with LEDS (0.81), LB_BHBA (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 mean absolute genetic correlation (MeanAbsCorr = 0.51), was identified as a key chromosome associated with LPNEB. The chromosome-specific MeanAbsCorr values for LPNEB are provided in Table A-6. The LPNEB on BTA25 showed the strongest genetic correlations with LEDS (0.88), LB_BHBA (0.74), and NEFA (0.87). Ha et al. (2015) found that genes related to BHBA 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 are supported by the known effect of NEB during lactation on milk fatty acid composition. Previous studies have reported associations between energy balance during early lactation in dairy cows and genomic regions 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.

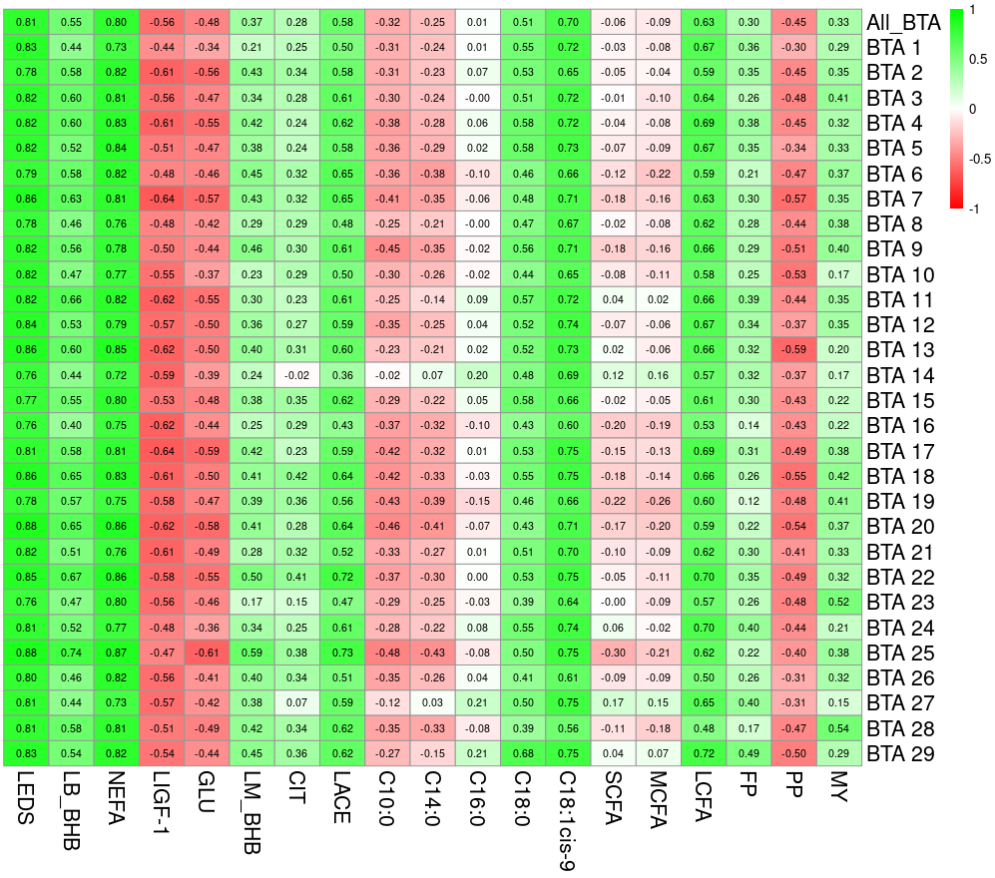


Figure 3-3. Genomic correlations between logarithm probability negative energy balance (LPNEB) and 19 other traits based on SNP effects. The first row represents genomic correlations estimated using SNP effects across all 29 chromosomes. Rows 2 to 30 represent genomic correlations estimated using SNP 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_BHBA= log_{10}-transformed blood beta-hydroxybutyrate predicted by MIR spectra; NEFA= blood non-esterified fatty acids predicted by MIR spectra; LIGF-1= log_{10}-transformed blood insulin-like growth factor 1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHBA = log_{10}-transformed milk beta-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.

3.4. Correlation between LEDS and 19 traits across different chromosomes

Figure 3-4 presents the genomic correlations between LEDS and 19 traits across different chromosomal regions. All significance tests of the genomic correlations presented in Figure 3-4 are reported in Table A-7. 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 mean absolute genetic correlation (MeanAbsCorr = 0.58), was identified as the key chromosome associated with LEDS. BTA25 showed the second highest MeanAbsCorr (0.57). The chromosome-specific MeanAbsCorr values for LEDS are provided in Table A-8. 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 C18:1cis-9 (0.84), LACE (0.84), and NEFA (0.82). Knutsen et al. (2022) reported that BTA19 is significantly associated with C18:1cis-9. These findings suggest that BTA19 may play a central role in the regulation of NEB in dairy cows during early lactation.

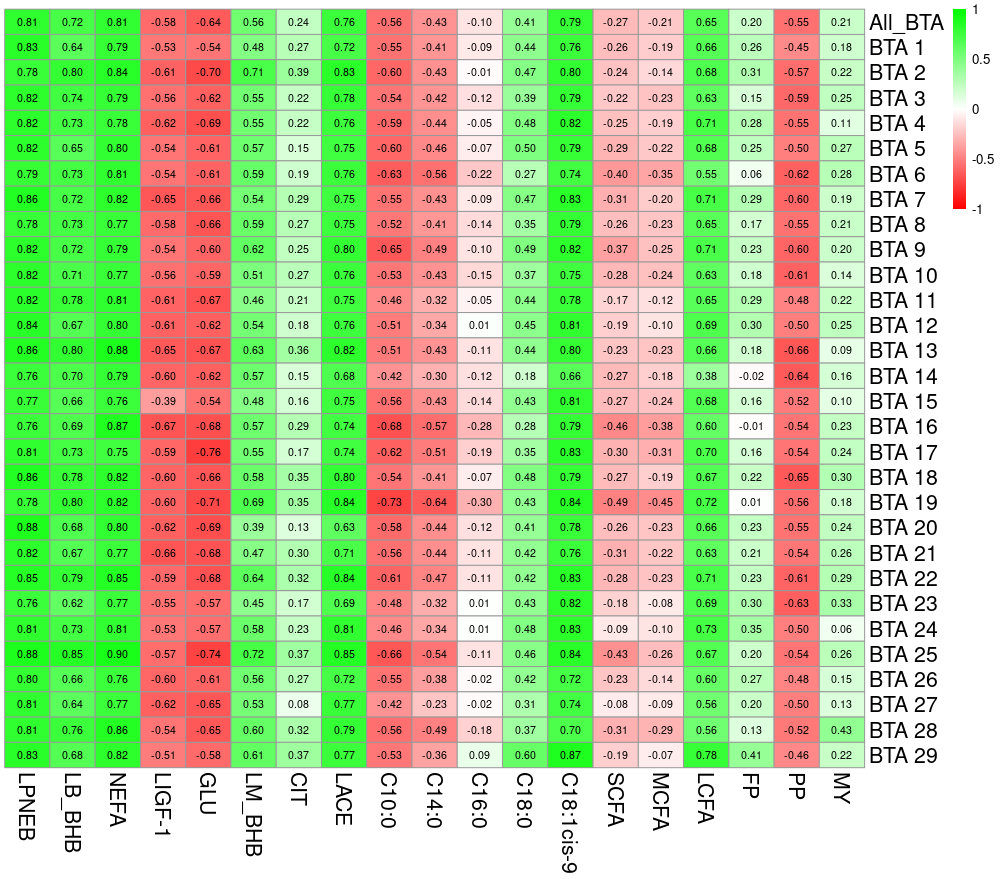


Figure 3-4. Genomic correlations between logarithm probability energy deficiency score (LEDS) and 19 other traits based on SNP effects. The first row represents genomic correlations estimated using SNP effects across all 29 chromosomes. Rows 2 to 30 represent genomic correlations estimated using SNP 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_BHBA= log_{10}-transformed blood beta-hydroxybutyrate predicted by MIR spectra; NEFA= blood non-esterified fatty acids predicted by MIR spectra; LIGF-1= log_{10}-transformed blood insulin-like growth factor 1 predicted by MIR spectra; GLU = blood glucose predicted by MIR spectra; LM_BHBA = log_{10}-transformed milk beta-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.

3.5. Independent contributions of 8 selected traits to LPNEB and LEDS

As shown in Figure 3-5, the selected 8 traits (NEFA, IGF-1, LB_BHBA, LM_BHBA, 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.

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 genomic 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 (Mansour et al., 2022; Mehtiö et al., 2020).

Although BHBA (including LB_BHBA and LM_BHBA) 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 fatty acid mobilization, BHBA likely reflects metabolic outcomes rather than acting as a direct genetic driver of NEB. 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, compared with the prediction accuracy of NEFA and LPNEB models, LEDS demonstrated superior predictive performance (accuracy = 0.99). Given that LEDS is both heritable and readily obtainable on a large scale from MIR spectra, we recommend its implementation as a proxy trait for NEB in dairy cattle breeding programs. However, 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 milk yield (MY), all phenotypes were predicted from MIR spectra rather than directly measured, which may artificially inflate genomic correlations due to shared prediction errors. 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 multi-trait models with individual-level data, such as multi-trait ssGREML, provide more accurate estimates, their application remains computationally challenging for large numbers of traits. Further methodological advances are needed to address this limitation.

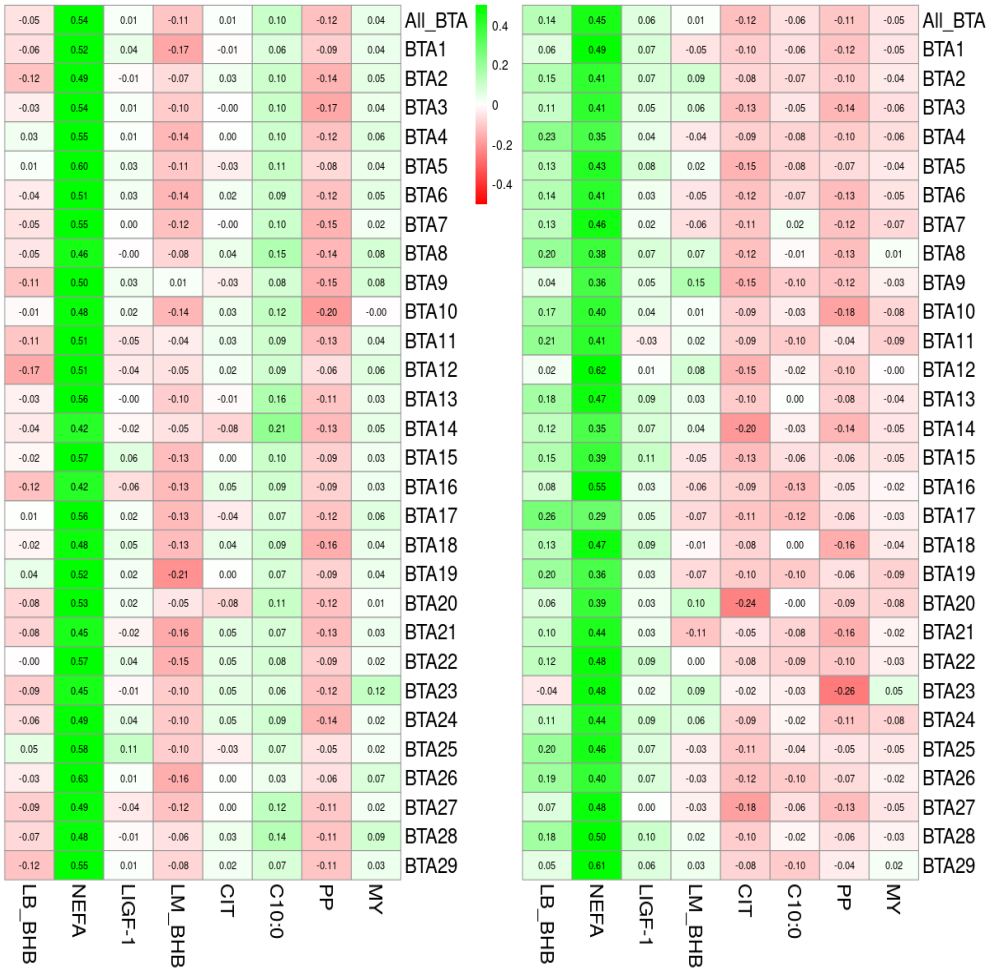


Figure 3-5. Independent contribution of 8 traits to LPNEB (left) and LEDS (right). The first row displays the independent contributions across all 29 chromosomes. Rows 2 to 30 represent independent contributions estimated from each individual chromosome. The 8 traits include: LB_BHBA= log_{10}-transformed blood beta-hydroxybutyrate predicted by MIR spectra; NEFA= blood non-esterified fatty acids predicted by MIR spectra; LIGF-1= log_{10}-transformed blood insulin-like growth factor 1 predicted by MIR spectra; LM_BHBA = log_{10}-transformed milk beta-hydroxybutyrate acid predicted by MIR spectra; CIT = milk citrate predicted by MIR spectra; C10:0 = milk decanoic acid predicted by MIR spectra; C18:1 cis-9 = milk oleic acid predicted by MIR spectra; FP = milk fat percentage predicted by MIR spectra; PP = milk protein percentage predicted by MIR spectra; MY = milk yield.

4. Conclusions

This study compared the genetic architecture of LPNEB and LEDS and their relationships with 15 biomarkers, 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. As SNP effects are computationally simpler and more suitable for large scale multi-trait genomic correlation analyses. Among all traits, NEFA showed strong genomic correlations to both LPNEB and LEDS, indicating its potential roles as direct genetic drivers of NEB. Chromosome level analyses identified key chromosomes such as BTA19 and 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.

5. Acknowledgments

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 provided also support through the PDR projects “HTwoTHI” (grant number T.W005.23) and “DEEPSELECT” (grant number T.0095.19). The authors declare that they have no competing interests. The Supplementary materials for this article are available 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.

Author contributions are as follows: Hongqing Hu: funding acquisition, formal analysis, writing – original draft; Sébastien Franceschini: data curation, writing – review and editing; Pauline Lemal: validation, writing – review and editing; Hadi Atashi: validation, writing – review and editing; Clément Grelet: data curation, writing – review and editing; Yansen Chen: data curation, writing – review and editing; Katrien Wijnrocx: validation, writing – review and editing; Soyeurt Hélène: Validation, writing – review and editing; Nicolas Gengler: conceptualization, funding acquisition, methodology, project administration, software, supervision, validation, writing – review and editing.

6. 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 April 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, 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.

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.

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., 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.

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.

Gardner, K.M., and R.G. Latta. 2007. Shared quantitative trait loci underlying the genetic correlation between continuous traits. Molecular Ecology 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., A. Vanlierde, F. Dehareng, E. Froidmont, and GplusE 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, Wageningen Academic Publishers. p. 403. https://hdl.handle.net/2268/224000.

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, 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, G. 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. Anim. 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. 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 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 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. 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 April 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 15:1–12. 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, 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. (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, Proceedings of the International Conference on Farm Animal Reproduction 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.2008-175.

4

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

Adapted from: Hu, H., H. Atashi, S. Franceschini, P. Lemal, C. Grelet, Y. Chen, K. Wijnrocx, H. Soyeurt, and N. Gengler. 2025. Using single-step genome-wide association analyses to compare predicted negative energy balance and a novel energy deficiency score in early lactation Holstein cows. JDS Communications. 6:792-796. https://doi.org/10.3168/jdsc.2025-0778.

Foreword

While the previous Chapter addressed the conventional indicator of energy status in dairy cows, this Chapter presents a comparative genetic analysis of MIR-predicted negative energy balance (LPNEB), a novel energy deficiency score (LEDS), and several MIR-predicted biomarkers. LPNEB and LEDS are closely related to energy metabolism and metabolic health, while the biomarkers provide complementary physiological information. Together, they may offer a more comprehensive and accurate representation of the early lactation metabolic status in dairy cows. These findings provide the first genetic references for LPNEB, LEDS, and associated biomarkers, supporting the development of breeding and management strategies to improve metabolic resilience, enhance production efficiency, and reduce the incidence of metabolic disorders.

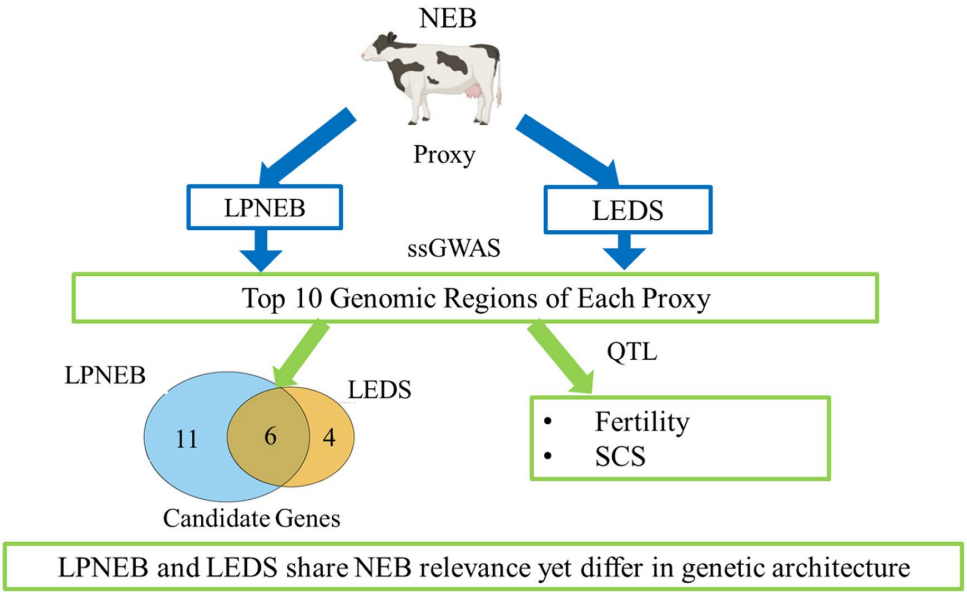


Figure 4-1. Genetic comparison of LPNEB and LEDS showing shared relevance to NEB but differences in genetic architecture.

Summary

This study compared the genetic architectures of 2 mid-infrared (MIR)-predicted proxies of negative energy balance (NEB) in early-lactation Holstein cows: the logit-transformed predicted NEB (LPNEB) and the logit-transformed energy deficiency score (LEDS). Using single-step genome-wide association study (ssGWAS) and functional enrichment analysis, the results revealed both overlapping and distinct genomic regions, biological pathways, and candidate genes for the 2 traits. 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.

Highlights

- Two MIR-predicted NEB proxies (LPNEB and LEDS) were compared in Holstein cows.
- Both LPNEB and LEDS exhibit polygenic architectures.
- The study provides new insights into the mechanisms underlying energy balance.

Abstract

This study aimed to compare the genetic architectures of logit-transformed predicted negative energy balance (LPNEB) and a novel logit-transformed energy deficiency score (LEDS) as 2 mid-infrared-derived proxies of negative energy balance (NEB) 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 SNP were available for 3,757 animals. Single-step genome-wide association analysis (ssGWAS), 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, while 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.

Key words: Negative energy balance, mid-infrared spectra, ssGWAS

1. Introduction

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 analysis 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., 2022). The EDS integrates multiple metabolic markers, making it a more reliable indicator of NEB than individual biomarkers. Cows in the EDS group show lower Body weight and DMI, along with higher BHBA 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 BHBA (0.69), and blood nonesterified 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, and (2) conducted comparative functional annotation of selected genomic regions.

2. Materials and methods

The overall study framework is presented in Figure 4-2.

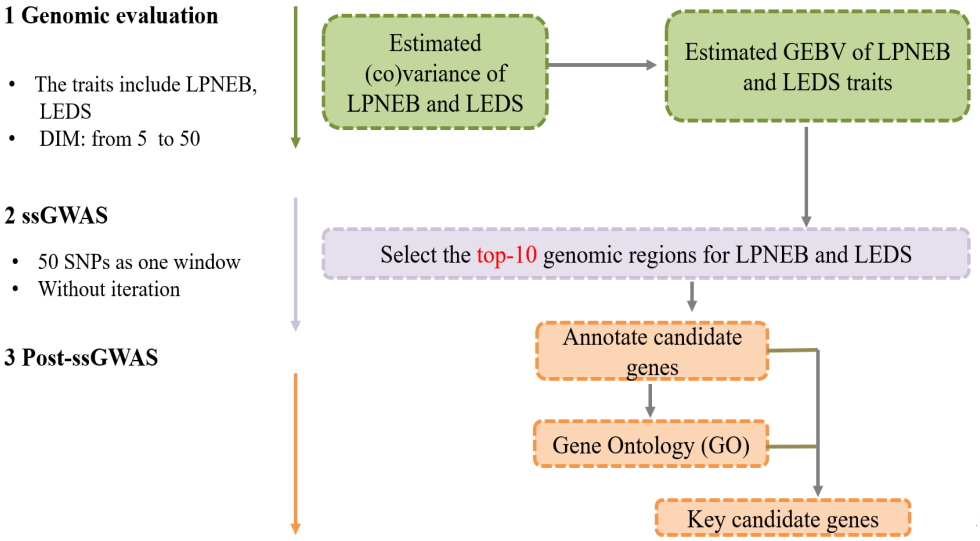


Figure 4-2. Workflow illustrating the data structure and analytical procedures used in this study.

2.1. Data

The data used in this study was 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 DIM 5 to 50. The MIR-based EB 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 SNP. 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:

$$\mathrm{LPNEB} = \log_{10}\frac{PNEB}{1 - PNEB} \quad (4.1)$$

where $\text{PNEB} = 1 - \frac{\text{PEB} - \text{PEB}_{\text{minimum}}}{\text{PEB}_{\text{maximun}} - \text{PEB}_{\text{minimum}}}$ (4.2)

$$\text{LEDS} = \log_{10} \frac{EDS}{1 - EDS} \quad (4.3)$$

2.2. (Co)variance components estimation

Similar to Hu et al. (2025), a univariate repeatability model was used to estimate variance components and genomic estimated breeding values (GEBV) for LPNEB and LEDS in first-parity cows. The model included fixed effects of herd × 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 non-linear effects of days in milk during early lactation. Random effects were the additive genetic effects, permanent environmental effects, and residual effects. The single-step genomic BLUP (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 defaults settings in BLUPF90+ software. Variance components were estimated using average information REML implemented in the BLUPF90+ software (version 2.48; Misztal et al., 2014). GEBV were computed using the precondition conjugate gradient algorithm implemented in BLUPF90+.

2.3. Genome-wide association analyses

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):

$$\hat{\mathbf{u}} = \mathbf{D}\mathbf{Z}_\mathbf{g}'\left[\mathbf{Z}_\mathbf{g}\mathbf{D}\,\mathbf{Z}_\mathbf{g}'\right]^{-1}\hat{\mathbf{a}} \quad (4.4)$$

where $\hat{\mathbf{u}}$ is the SNP effect, D is the weight matrix of SNPs (**D** = **I**), which means the weight for all SNPs is 1; **Zg** is an incidence matrix of genotypes for each SNP; and $\hat{\mathbf{a}}$ is a vector of genomic EBV for each trait of genotyped animals. The variance of ith SNP is following: $var(\hat{u}_i) = u_i^2 \cdot 2p_i(1 - p_i)$, where u_i^2 is the square of ith SNP effect, and pi 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 about 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:

$$\frac{\mathrm{var}(\hat{\mathbf{a}}_i)}{{\sigma_a}^2} \times 100 = \frac{\mathrm{var}(\boldsymbol{Z_i}\hat{\mathbf{u}}_i)}{{\boldsymbol{\sigma_a}}^2} \times 100 \quad (4.5)$$

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, and $\hat{\mathbf{u}}_i$ the vector of SNP all defined in the ith genomic region and ${\sigma_a}^2$ is the total additive genetic variance. We used a window of 50 consecutive SNPs following Atashi et al. (2024), as this size provides a suitable balance between mapping resolution and the coverage of a meaningful genomic region.

2.4. Functional annotation analyses

Genetic 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, accessed on Feb. 02, 2025; 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). 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 (4.2.1; R Foundation for Statistical Computing).

3. Results and discussion

3.1. Genome-wide association analyses

Figure 4-3 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 4-1 compares the top 10 genomic regions explaining the biggest 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. BTA 5 and 16 were identified in genomic regions with significant effects on energy balance (Tetens et al., 2013; Krattenmacher et al., 2019). Additionally, unique genomic regions are associated with each trait, with LPNEB showing specific regions on BTA 6, 13, and 25, while LEDS exhibits distinct regions on BTA 7, 9, and 17. These differences indicate potential variations in the genetic architecture of the two 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 two 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.

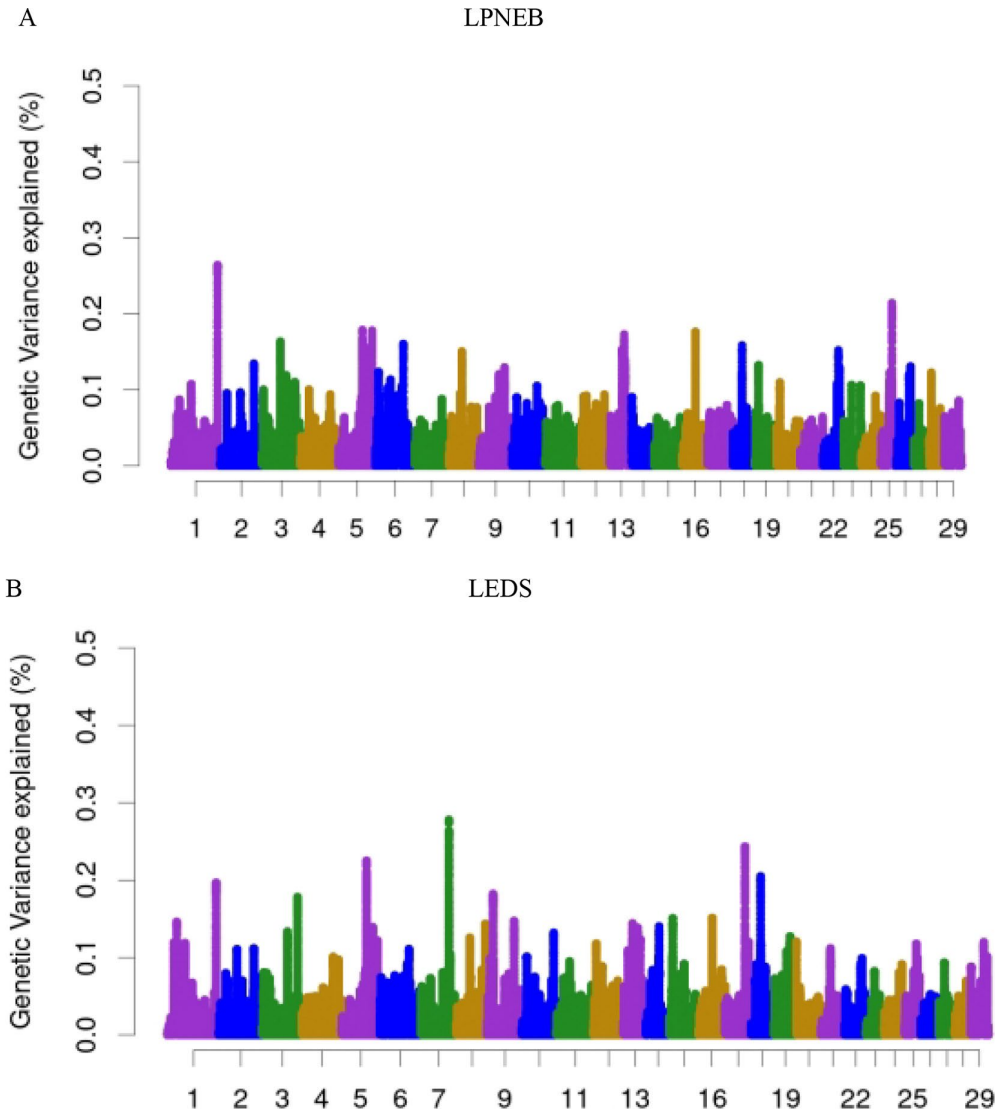


Figure 4-3. Additive genetic variance explained by windows of 50 adjacent SNPs across chromosomes for logit-transformed predicted negative energy balance (LPNEB, A) and logit-transformed energy deficiency score (LEDS, B) in early-lactation Holstein cows.

3.2. Gene annotation analyses

The genes annotation of the selected top 10 genomic regions for LPNEB and LEDS identified a total of 17 positional candidate 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 notable overlap in metabolic processes (Figure 4-4A and 4-4B), particularly in D-gluconate metabolic (GO: 0019521), D-gluconate catabolic (GO: 0046177) and pentose biosynthesis (GO: 0019322), all influenced by *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, while LEDS integrates neuronal signaling into energy homeostasis.

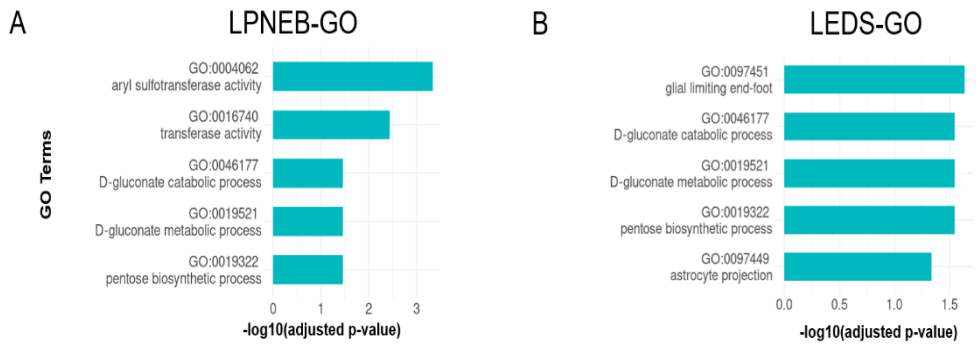


Figure 4-4. The Gene Ontology (GO) terms (A and B) based on candidate genes of logit-transformed predicted negative energy balance (LPNEB) and logit-transformed energy deficit score (LEDS).

Table 4-1. Annotated genes within the top ten 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).

Trait	BTA[1]	Position(bp)[2]	Var[3]	Gene[4]
LPNEB	**1**	**150974434:151093598[5]**	**0.26**	***PIGP, TTC3***
	3	55153255:55383999	0.16	*KYAT3, GTF2B, PKN2*
	5	**82540690:82638650**	0.18	***PPFIBP1***
	5	101175264:101308905	0.16	
	5	111857928:112120252	0.18	*FAM83F, TNRC6B*
	6	86686814:87001138	0.16	*ENSBTAG00000015047, UGT2A1, SULT1B1, ENSBTAG00000038214*
	13	53637893:53759622	0.17	*SIRPA*
	16	**44046234:44165527**	**0.18**	***PGD, ENSBTAG00000048790, ENSBTAG00000048747, ENSBTAG00000054239***
	18	25753024:25847046	0.16	
	25	30079410:30209679	0.21	
LEDS	**1**	**151023504:151147932**	**0.20**	***TTC3,***
	3	109338317:109464393	0.18	*GRIK3*
	5	**82540690:82638650**	**0.23**	***PPFIBP1***
	7	88653706:88732303	0.28	
	9	18311181:18423364	0.18	
	9	84820057:84931668	0.15	
	15	10352544:10716253	0.15	
	16	**44037583:44158583**	**0.15**	***PGD, ENSBTAG00000048790 ENSBTAG00000048747, ENSBTAG00000054239***
	17	58683435:58772056	0.24	
	18	25580320:25699738	0.21	*CCDC102A, ADGRG5, ADGRG1*

[1] BTA = Bos taurus autosomes.

[2] Position(bp) = Starting and ending coordinates of the genomic region.

[3] Var = percentage of genetic variance explained by the genomic regions within the genomic region.

[4]Genes = EBSEMBL symbol of annotated genes using the Bos Taurus ARS-UCD 1.2 assembly (http://ftp.ensembl.org/pub/release-110/gtf/bos_taurus/).

[5]Overlapping genomic regions between LPNEB and LEDS are highlighted in bold.

3.3. QTL annotation for select genomic regions

The QTL enrichment analysis of the top 10 genomic regions associated with LPNEB and LEDS against the cattle QTLdb identified substantial associations with multiple economically important traits (Figure 4-5C and 4-5D). QTL enrichment analyses of LPNEB and LEDS revealed two shared enriched regions: somatic cell score (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 that 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.

Figure 4-5. The QTL annotation (C and D) based on candidate genes of logit-transformed predicted negative energy balance (LPNEB) and logit-transformed energy deficit score (LEDS).

4. Conclusions

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 LNEB 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.

5. Acknowledgments

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 Projet de Recherche projects “HTwoTHI” (grant number T.W005.23) and “DEEPSELECT” (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 declare that they have no competing interests.

6. References

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

Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. PREGSF90–POSTGSF90: Computational tools for the implementation of single-step genomic selection and genome-wide association with ungenotyped individuals in BLUPF90 programs. 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.

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.

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, Ma. L, T. S. Sonstegard, T. J. Lawlor, B. A. Crooker, C. Tassell, J. Yang, S. Wang, L. Matukumalli, and Da. Y. 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:giaa149. https://doi.org/10.1093/gigascience/giaa149.

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 GplusE 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, Wageningen Academic Publishers. p. 403. https://hdl.handle.net/2268/224000.

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. In press. 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.

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.

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 Research, Volume 51:W207–W212, https://doi.org/10.1093/nar/gkad347.

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 economically 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. 03, 2025. 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 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.

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. Animal genetics, 44(1): 1-8. https://doi.org/10.1111/j.1365-2052.2012.02357.x.

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.

van den Berg, I., R. Xiang, J. Jenko, H. Pausch, M. Boussaha, C. Schrooten, T.Tribout, A. Gjuvsland, 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. GSE 52: 37. https://doi.org/10.1186/s12711-020-00556-4.

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.

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. (Camb.) 94:73–83. https://doi.org/10.1017/S0016672312000274.

Zachut, M., M. Šperanda, A. M. de Almeida, G. Gabai, A. Mobasheri, 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.

5

Chapter V General discussion, conclusions and perspectives

Foreword

To close this thesis, a synthesis is drawn across Chapters 2 to 4, each addressing complementary aspects of the genetic architecture of negative energy balance (NEB) in early-lactation dairy cows.

***Chapter 2** focused on the comparison relationships of MIR-predicted NEB. Predicted NEB (LPNEB) was validated against measured reference values, and a novel energy deficiency score (LEDS) was introduced as a composite proxy with greater explanatory power.*

***Chapter 3** applied SNP-based genomic correlation and independent contribution analyses, confirming that LEDS and LPNEB exhibit similar genetic relationships with metabolic biomarkers, with NEFA emerging as the central genetic driver of NEB and BTA19 and BTA25 identified as major shared genomic regions associated with these biomarkers.*

***Chapter 4** investigated the genomic architecture of LPNEB and LEDS through single-step GWAS, revealing their polygenic nature, key functional pathways, and overlaps with fertility, immunity, and production traits.*

Finally, the thesis suggests that MIR-derived proxies, particularly LEDS, can be incorporated into genomic selection and breeding strategies to enhance metabolic resilience in dairy cattle.

1. General discussion

This thesis provides an integrated understanding of the genetic architecture underlying negative energy balance (NEB) in early-lactation Holstein cows by combining pedigree-based analyses, SNP based genomic correlations, Genome-wide association studies (ssGWAS), and genomic validation. The results consistently demonstrate that both MIR-predicted indicators logit-transformed predicted NEB (LPNEB) and the composite logit-transformed novel energy deficiency score (LEDS) capture the genetic basis of NEB with high concordance (rg > 0.8), yet reflect partially distinct biological processes. nonesterified fatty acids (NEFA) emerged as the most informative and causally relevant biomarker, confirming its central role in fat mobilization during early lactation. Chromosome level analyses identified BTA19 and BTA25 as key genomic regions regulating NEB and associated lipid metabolism. While both LPNEB and LEDS exhibited highly polygenic inheritance, ssGWAS revealed overlapping but distinct genomic regions, with LPNEB being more tightly linked to lactation metabolism and LEDS encompassing broader pathways related to immunity and physiological resilience. Validation analyses further showed that LEDS displayed more robust cross-population predictive ability than LPNEB, likely due to its composite design that integrates multiple mid-infrared spectrometry (MIR) predicted biomarkers and mitigates environmental variability. Overall, this study provides deeper insights into the genetic basis of NEB and the genetic interrelationships among associated metabolic traits, highlighting LEDS as a promising high throughput phenotyping and genomic selection indicator for enhancing metabolic efficiency and physiological resilience in dairy cattle.

1.1. Genetic correlations between LPNEB, LEDS, and biomarkers

To investigate the genetic correlations of NEB and its related traits, we first estimated pedigree based genetic correlations using a 20-trait repeatability animal model. This framework allowed a simultaneous evaluation of LPNEB, LEDS, and 18 traits, including MIR-predicted biomarkers and directly measured production traits such as milk yield (MY). The strong genetic correlation between LPNEB and LEDS (rg = 0.85) indicates that these two traits share a close genetic correlation in the context of energy metabolism (Chapter 2, Figures 2–5; Hu et al., 2025). Among all traits, NEFA displayed the strongest genetic correlation with LPNEB (rg = 0.87), which reflects their shared biological role in fat mobilization during early lactation (Mehtiö et al., 2020; Mansour et al., 2022). Similarly, milk fatty acids, particularly milk oleic fatty acid (C18:1 cis-9) and milk long-chain fatty acids (LCFA), showed strong genetic correlations with LPNEB (rg > 0.60), consistent with their established

involvement in lipid mobilization and β-oxidation (Bastin et al., 2012; Churakov et al., 2021; Martens, 2023).

Because the 20-trait repeatability model is computationally demanding, we further incorporated genomic data by first estimating SNP effects for each trait using single-trait repeatability models, and then calculating SNP-based genomic correlations among the 20 traits. As shown in Chapter 3 (Figure 3-2), the genomic correlations were estimated using single-trait repeatability models based on SNP effects. These genomic correlations were generally consistent with pedigree-based genetic correlations. The pedigree-based correlations were obtained from the 20-trait repeatability model. This consistency indicates that SNP-based genomic correlations can effectively reproduce and compute the genetic correlation results of the 20-trait repeatability model, while providing a simpler and more efficient computational alternative. Importantly, this strategy has rarely been systematically applied in previous studies, thus representing a key methodological innovation of the present work and offering a novel framework for large-scale evaluation of genetic relationships among complex traits. For example, LPNEB and LEDS showed a very strong genomic correlation (0.81), while NEFA again maintained the highest genomic correlations with both LPNEB (0.80) and LEDS (0.81). Beta-hydroxybutyrate acid (BHBA) and acetone showed moderate genomic correlations, consistent with their roles as ketone bodies produced during fat mobilization (Guliński, 2021; Lisuzzo et al., 2022). Likewise, the genomic correlations of C18:1 cis-9 with LPNEB (0.70) and LEDS (0.79) are in line with previous reports showing that milk fatty acid composition can serve as a reliable proxy of NEB (Knutsen et al., 2022; Xu et al., 2020). Overall, genetic correlations derived from SNP effects based on single trait models provides a simpler and more computationally efficient alternative to conventional multi-trait approaches. This strategy markedly reduces computational demands while still capturing the essential correlation structure among traits, making it particularly suitable for studies involving large numbers of traits.

To further refine the analysis, we calculated chromosome level genomic correlations based on SNP effects. This approach enabled finer dissection of shared genetic architecture between NEB traits and related biomarkers. For LPNEB, BTA25 emerged as the most influential chromosome (MeanAbsCorr = 0.51), showing strong genomic correlations with NEFA (0.87) and C18:1 cis-9 (0.75). For LEDS, BTA19 was the most prominent chromosome (MeanAbsCorr = 0.58), with very strong genomic correlations with NEFA (0.82), C18:1 cis-9 (0.84), and acetone (0.84). These results are consistent with previous studies reporting QTL for lipid metabolism and milk fatty acids on BTA19 and BTA25 (Tetens et al., 2013; Knutsen et al., 2022). These differences also indicate that LEDS and LPNEB do not rely on identical genomic signals. Because LEDS integrates multiple MIR-predicted biomarkers, it

captures a broader metabolic signature and therefore aligns more strongly with loci on BTA19 that regulate lipid mobilization and fatty-acid metabolism. In contrast, LPNEB is derived directly from MIR-predicted NEB and reflects spectral features that appear to be more strongly influenced by variants on BTA25. This difference in trait definition likely explains why the strongest chromosome-level associations of LEDS occur on BTA19, whereas LPNEB shows its peak contribution on BTA25.

Collectively, the integration of pedigree-based genetic correlations estimated from a multi trait repeatability model, genomic correlations derived from single trait models based on SNP effects, and chromosome level genomic correlations provides a consistent and biologically meaningful understanding of NEB and its related traits. Pedigree based analyses capture the overall interrelationships among different traits, whereas SNP effect single-trait models not only reproduce these correlation patterns but also offer a more computationally efficient framework. Furthermore, chromosome-level genomic correlations enable the fine mapping of shared genomic regions.

The results indicate that correlations estimated from both pedigree and SNP effect-based approaches were highly consistent, with the correlation between LPNEB and LEDS exceeding 0.8, suggesting that the two traits largely capture overlapping genetic signals. Across all levels of analysis, NEFA consistently emerged as the most informative biomarker, while BTA19 and BTA25 were highlighted as key genomic regions influencing the genetic regulation of NEB during early lactation. These findings not only extend pedigree based analyses (Hu et al., 2025) but also align with genomic studies of fat mobilization and energy metabolism (Tetens et al., 2013; Krattenmacher et al., 2019; Knutsen et al., 2022).

1.2. Causal effects and independent contributions between LPNEB, LEDS, and biomarkers

Based on the 20-trait repeatability model described in Chapter 2, we further applied a recursive model to explore causal effects between LPNEB and other traits. The results showed that NEFA, LEDS, and C18:1 cis-9 exerted the strongest causal effects on LPNEB, which is consistent with their central roles in energy metabolism (Mehtiö et al., 2020; Mansour et al., 2022; Churakov et al., 2021). NEFA, in particular, showed the largest structural coefficients and variance contributions, confirming its pivotal role in fat mobilization during early lactation. LEDS, which integrates multiple biomarkers such as BHBA, milk citrate (CIT), and C18:1 cis-9, also exhibited significant causal effects on LPNEB, reinforcing its value as a composite indicator (Hu et al., 2025). In contrast, milk yield showed an opposite causal direction compared with its genetic correlation, suggesting that higher milk production may drive NEB

rather than being a direct outcome of NEB, in line with the energy partitioning theory in early lactation (Martens, 2023).

We further evaluated the independent contributions of 8 key biomarkers (NEFA, log10 blood insulin-like growth factor 1 (LIGF-1), log10 blood beta-hydroxybutyrate acid (LB_BHBA), log10 milk beta-hydroxybutyrate acid (LM_BHBA), CIT, milk decanoic fatty acid (C10:0), milk protein percentage (PP), and MY) to LPNEB and LEDS using SNP effects. The results revealed that NEFA consistently accounted for the largest independent contribution, maintaining strong genomic correlations with both LPNEB and LEDS, while also exerting a significant direct effect after accounting for other traits (Mehtiö et al., 2020). By contrast, BHB (LB_BHBA and LM_BHBA), although moderately correlated with NEB traits, showed relatively low independent contributions, suggesting that ketone bodies mainly reflect downstream metabolic consequences of fat mobilization rather than acting as direct genetic drivers (Guliński, 2021; Lisuzzo et al., 2022). Similarly, LIGF-1 had a low independent contribution, indicating that its role is likely indirect through endocrine regulation (Butler, 2003). Interestingly, protein percentage (PP) exhibited consistent negative contributions, suggesting a potential negative genetic correlation with NEB, which is consistent with the competition between energy allocation and milk protein synthesis (Bastin et al., 2012). Other traits, such as C10:0, CIT, and MY, made only minor independent contributions, indicating a limited direct genetic role in NEB regulation.

In summary, the combined evidence from the recursive model and independent contribution analyses consistently demonstrated that NEFA is the most critical genetic biomarker of NEB, highlighting its central role in fat mobilization. Meanwhile, LEDS showed clear advantages as a composite indicator by integrating multiple metabolic signals and capturing broader aspects of energy deficiency. These findings deepen our understanding of the causal relationships and independent contributions of biomarkers to NEB.

1.3. Genetic architecture of LPNEB and LEDS revealed by ssGWAS

Genome-wide association studies (ssGWAS) confirmed that both LPNEB and LEDS follow a highly polygenic inheritance pattern, with no major loci detected. The top 10 windows explained less than 0.5% of the variance, emphasizing the dispersed genetic control typical of complex metabolic traits. This polygenic background aligns with previous studies on energy balance in dairy cattle (Buitenhuis et al., 2016; van den Berg et al., 2020; Pedrosa et al., 2021).

Despite this overall polygenic nature, several chromosomes were consistently implicated in both LPNEB and LEDS. including BTA1, BTA5, and BTA16, suggesting overlapping genetic control of energy metabolism. Candidate gene

annotation further highlighted loci such as *PGD, TTC3*, and *PPFIBP1*, which are involved in energy metabolism, amino acid turnover, and transcriptional regulation. Functional enrichment analyses revealed that LPNEB was more strongly associated with metabolic detoxification and enzymatic transferase activity, while LEDS encompassed a broader set of biological processes, including neuronal signaling and immune pathways, reflecting its design as a composite trait.

To strengthen the biological interpretation of these findings, future work should incorporate gene expression evidence from relevant tissues, such as liver, mammary gland, and adipose tissue, to validate whether the identified candidate genes show differential expression across cows with contrasting NEB status. Integrating GWAS signals with transcriptomic resources (e.g., RNA-seq or publicly available cattle expression atlases) would allow the identification of expression quantitative trait loci (eQTLs) and help establish whether these variants regulate transcriptional activity in metabolic pathways. Such integrative analyses could provide stronger causal support for the involvement of these genes in early-lactation energy metabolism and offer promising targets for functional validation and breeding applications.

Distinct chromosomal regions also emerged: LPNEB exhibited strong associations on BTA6, BTA13, and BTA25, consistent with regions previously linked to milk yield, milk protein percentage, and lipid mobilization (Ha et al., 2015). LEDS, on the other hand, was associated with BTA7, BTA9, and BTA17, enriched for QTL related to immunity, muscle development, and growth. For instance, QTL enrichment for muscle zinc content on BTA9 highlights the role of zinc in protein synthesis and immune function (Hawken et al., 2012; Mateescu et al., 2017). These results demonstrate that while LPNEB is more directly tied to lactation-related pathways, LEDS integrates additional biological dimensions, such as immunity and growth, underscoring its broader utility for resilience traits.

1.4. Validation of genomic predictive accuracy for LPNEB and LEDS?

Genomic predictive ability of NEB was assessed using SNP effects of LPNEB and LEDS estimated in Chapter 3. The evaluation was based on 92 cows with logarithm meanured NEB (LRNEB) values and available genotypes. Direct genomic values (DGV) were obtained by combining SNP effects with genomic data. LRNEB phenotypes were adjusted using a linear model with country, DIM, and parity as fixed effects and animal as a random effect. Predictive ability was quantified as the Pearson correlation between DGVs of LPNEB or LEDS and adjusted LRNEB phenotypes.

The validation analyses demonstrated that the genomic predictive ability of LPNEB and LEDS for measured NEB (LRNEB) was modest, The Pearson correlation between the DGV of LPNEB and adjusted LRNEB was 0.10, while that between the

DGV of LEDS and adjusted LRNEB was 0.13. Such results are not unexpected for complex metabolic traits like NEB, which are strongly influenced by environmental and management factors.

A major explanation for the reduced predictive ability lies in differences between reference and validation populations. LPNEB predictions were derived mainly from Belgian primiparous Holsteins, whereas LRNEB reference values originated from Danish Holsteins, often across multiple parities (Aarhus University). Discrepancies in genetic background, management systems, and physiological stage likely contributed to reduced consistency, consistent with earlier findings that predictive accuracies decrease when populations differ in origin or production environment (McParland et al., 2015; van den Berg et al., 2020). Feeding practices also represent a key source of variation: Belgian herds are commonly managed under pasture-based or mixed systems, while Danish herds typically use intensive housing with distinct forage-to-concentrate ratios. Such nutritional differences directly affect dry matter intake, milk yield, and energy partitioning, thereby contributing to discrepancies between predicted LPNEB and measured LRNEB. Furthermore, NEB is a dynamic trait that varies considerably during early lactation, and inconsistencies in measurement timing across herds may have further reduced predictive accuracy.

Interestingly, LEDS showed higher predictive consistency across populations compared with LPNEB. This likely reflects its composite design, which integrates multiple MIR-predicted biomarkers and thus captures a broader set of genetic signals while being less sensitive to environmental variation and calibration differences. By contrast, LPNEB, as a direct proxy of NEB, appears more vulnerable to errors introduced by spectral calibration and feeding system discrepancies. These results align with previous reports that composite or multi-trait indicators are more robust across populations (Franceschini et al., 2022; Grelet et al., 2021).

From a breeding perspective, although the predictive ability of both traits remains limited, LEDS appears to be the more promising candidate for genomic selection. To improve prediction accuracies, future studies should build larger multi-country reference populations, while also constructing validation populations that are more consistent with the reference, ideally based on primiparous cows from similar environments. Incorporating detailed feeding and management data, along with applying multi-trait genomic models, may further enhance the robustness of genomic evaluations for NEB.

2. Conclusions

This thesis provided an integrated understanding of NEB in first-parity Holstein cows during early lactation by combining MIR-derived phenotypes with pedigree and genome based analyses. In accordance with the defined research objectives, the work

systematically explored (1) the phenotypic and genetic correlations between MIR-predicted NEB indicators and multiple metabolic and production biomarkers, and (2) the genomic architecture underlying NEB and its MIR-derived traits. The main findings and implications are summarized as follows.

First, by integrating 20 traits within a repeatability model, strong phenotypic and genetic correlations were identified between the MIR-predicted indicators (LPNEB and LEDS) and key metabolic biomarkers. NEFA consistently emerged as the most informative and causally relevant biomarker, displaying the highest genetic correlations with both LPNEB and LEDS, followed by milk fatty acids such as C18:1 cis-9 and LCFA. LPNEB effectively captured the core metabolic dynamics of energy balance, while LEDS constructed from 27 MIR-predicted metabolic traits demonstrated superior predictive performance (accuracy = 0.99), confirming its robustness and scalability compared with single biomarker predictions such as NEFA. Second, SNP-based genomic correlation and independent contribution analyses revealed overlapping yet distinct genetic architectures between LPNEB, LEDS, and their associated biomarkers. NEFA showed the largest independent contribution to both indicators, reaffirming its central role in fat mobilization and lipid metabolism, whereas BHBA and IGF-1 showed more limited effects, indicating their secondary involvement in NEB regulation. Third, the ssGWAS results identified both shared and trait specific genomic regions underlying LPNEB and LEDS. LPNEB exhibited its strongest associations on BTA25 with NEFA, BHBA, and C18:1 cis-9, while LEDS was mainly associated with NEFA and milk fatty acids on BTA19. Shared genomic signals across BTA1, BTA5, and BTA16 indicated common regulatory pathways related to energy metabolism, whereas trait specific loci reflected distinct biological mechanisms. Candidate genes such as *PGD*, *TTC3*, and *PPFIBP1* were highlighted for their potential involvement in lipid and energy metabolism.

Overall, this work demonstrates that integrating MIR-predicted phenotypes with genomic analyses offers a powerful framework for dissecting the biological and genetic basis of NEB. Both LPNEB and LEDS capture the underlying genetic regulation of energy balance, yet LEDS, as a composite and highly accurate indicator, provides greater robustness and transferability across environments. These results support the implementation of MIR-based indicators particularly LEDS as reliable, scalable tools for genetic evaluation, metabolic monitoring, and selection toward improved resilience and energy efficiency in dairy cows.

3. Perspectives

3.1. Validation of MIR-derived NEB indicators using reference NEB and biomarkers

Future research on NEB should place strong emphasis on validation of MIR-derived proxies such as LPNEB and LEDS against reference NEB values, which remain the biological gold standard. Such validation should be performed in well-characterized Belgian dairy populations, and within the same parity, to avoid confounding effects of lactation stage or parity order. By selecting cows with high and low predicted values and recording their reference NEB, the predictive ability and robustness of MIR-based indicators can be directly assessed. In addition, it will be important to measure a broader panel of biological indicators to complement existing markers such as NEFA, BHBA, IGF-1, and milk fatty acids. Measuring multiple biomarkers alongside reference NEB will help to clarify which indicators are primary drivers of NEB (e.g., NEFA, specific fatty acids) and which represent downstream or secondary responses (e.g., BHBA, IGF-1). Such comparisons will not only provide deeper biological insight into energy balance regulation, but also guide the refinement of MIR-derived proxies by identifying the most relevant biological signals.

3.2. Broadening NEB research from primiparous to multiparous cows

This study primarily focused on primiparous cows, which are particularly relevant for early monitoring and genetic evaluation due to the availability of large-scale data. However, as parity increases, the energy requirements and metabolic stress of dairy cows rise substantially. In later lactations, especially at peak milk yield, multiparous cows must not only sustain higher milk production but also cope with cumulative energy expenditure, body condition recovery, and reproductive challenges, making NEB more pronounced and physiologically complex (Butler, 2003; Wathes et al., 2007; Ingvartsen and Andersen, 2000). Changes can be expected in later parities because metabolic demands, hormonal regulation, and nutrient partitioning patterns differ substantially from those in primiparous cows (Wathes et al., 2007).

Future research should therefore extend the analyses of LPNEB, LEDS, and associated biomarkers to multiparous cows to capture parity-specific patterns of genetic variation. At the phenotypic level, Czaplicki (2022) examined the transferability of the novel EDS trait from first to later lactations, but the mixed outcomes do not necessarily reflect the underlying genetic relationships. Despite this limitation, exploring this extension remains essential for addressing a key question: whether the same genomic regions and biomarkers consistently drive NEB across different parities, or whether distinct metabolic pathways and genetic mechanisms become predominant in later lactations. For example, NEFA and milk fatty acids have

been confirmed as strongly associated with NEB in primiparous cows (Song et al., 2021; Santos et al., 2001), but whether their contributions change in multiparous cows remains to be verified.

In addition, investigating parity specific effects will not only enhance our understanding of the complexity of NEB from a genetic perspective but also provide practical guidance for herd management and breeding. Identifying unique genetic signals and metabolic biomarkers in multiparous cows could inform the development of parity specific management strategies and breeding goals. This would ultimately improve lifetime resilience, reduce culling rates, extend productive longevity, and enhance overall herd productivity (Pinedo et al., 2014; Manafiazar et al., 2016).

Therefore, extending the research focus from primiparous to multiparous cows will allow for a more comprehensive understanding of the genetic architecture and physiological mechanisms of NEB, providing a solid foundation for targeted nutritional interventions and refined breeding objectives.

3.3. Application of MIR-derived NEB indicators in farm management and breeding

Using MIR-derived NEB indicators requires easy access to standardized MIR spectra. However, such access is still limited in many countries, including China. Once available, LPNEB and LEDS could be integrated into practical dairy production and breeding programs. These two MIR-derived proxies capture complementary aspects of NEB: LPNEB shows moderate correlations with reference NEB, reflecting key biological signals of energy metabolism, whereas LEDS, constructed from 27 MIR-predicted metabolic traits including milk BHBA and C18:1 cis-9, has demonstrated outstanding predictive performance. The prediction model, developed using partial least squares discriminant analysis, achieved an overall accuracy of 0.99 with high sensitivity (0.95) and specificity (0.92), clearly outperforming MIR-based NEFA prediction ($R^2cv = 0.39$) (Franceschini et al., 2022; Franceschini et al., 2024; Grelet et al., 2019). These features highlight LEDS as a more robust and scalable proxy, suitable for low-cost, high-throughput phenotyping in routine dairy herd management.

Combining LPNEB and LEDS can provide multidimensional monitoring of energy status at the farm level. This integrated approach may enable early detection of high-risk cows, thereby reducing the incidence of metabolic disorders such as ketosis and fatty liver (McArt et al., 2012; Ospina et al., 2010), while also informing precision nutrition and lactation management strategies (Ingvartsen and Andersen, 2000; Allen et al., 2009). From a breeding perspective, incorporating both LPNEB and LEDS into multi-trait genomic selection models will allow better exploitation of genetic correlations, supporting simultaneous improvement of metabolic resilience, reproductive efficiency, and productivity (Connor, 2015; van den Berg et al., 2020).

Moreover, integrating MIR-derived NEB proxies with traditional health and fertility traits could help balance the long-term trade-off between production and robustness, promoting sustainable dairy breeding (Butler and Smith, 1989; Brito et al., 2021).

Ultimately, translating LPNEB and LEDS into practical farm tools offers strong potential to advance precision livestock farming. Their high predictive power, biological relevance, and scalability make them valuable assets for herd-level monitoring, nutritional management, and genetic improvement, paving the way for more resilient and sustainable dairy production systems.

The usefulness of MIR-derived NEB indicators, especially for farm management, also depends on frequent sampling to obtain MIR spectra. In Wallonia, milk recording samples are collected every four or even six weeks, which is obviously suboptimal for management purposes. Different efforts are ongoing to implement adapted visible and near-infrared spectroscopic technologies on farm (Aernouts et al., 2011). If NEB indicators could be derived from this type of spectral data, their practical usefulness would be greatly increased.

3.4. Integrating physiological and management factors into NEB evaluations

Although this thesis mainly explored the genetic relationships of NEB with MIR-predicted traits and metabolic biomarkers, NEB is in reality a multifactorial condition influenced by a wide range of physiological, metabolic, and environmental factors. Its occurrence and severity are not only determined by feed intake and milk yield but are also strongly affected by parity, body condition at calving, hormonal regulation, liver metabolic capacity, health status, and environmental and management conditions. At the physiological level, BCS at calving is a major determinant of NEB dynamics. Cows with excessively high BCS are more prone to excessive fat mobilization, increasing the risk of ketosis and fatty liver, while cows with low BCS may lack sufficient energy reserves to support milk production and reproductive recovery (Roche et al., 2006). Parity also plays a crucial role: primiparous cows experience relatively moderate metabolic stress, whereas multiparous cows, particularly at peak lactation, must cope with both higher milk yields and cumulative reproductive burdens, resulting in more severe and prolonged NEB (Butler, 2003; Wathes et al., 2007).

In terms of metabolic regulation, hormonal status and liver capacity are central factors shaping NEB. Insulin, glucagon, growth hormone, and IGF-1 orchestrate glucose metabolism, lipolysis, and hepatic ketogenesis (Lucy, 2001). Individual variation in hepatic capacity for NEFA oxidation and very low density lipoprotein (VLDL) synthesis contributes to differences in energy utilization efficiency and susceptibility to metabolic disorders (Drackley, 1999). At the environmental and

management level, feeding systems, diet composition, heat stress, milking frequency, and grouping strategies all influence NEB. For example, under heat stress, cows reduce dry matter intake while energy demands increase due to thermoregulation, thereby exacerbating NEB (Rhoads et al., 2009). Similarly, peripartum nutrition, grouping strategies, and herd health management can either mitigate or exacerbate NEB risk.

Future research should integrate these broader factors into genomic evaluations. Building multidimensional datasets that combine BCS, health records, environmental variables, and management information with MIR-derived NEB indicators would allow better disentangling of genetic and environmental effects. Moreover, applying multi-trait or reaction norm models can capture genotype-by-environment interactions and trait variability under diverse conditions (Kolmodin and Bijma, 2004). Ultimately, linking NEB indicators with these multifactorial influences will provide a more holistic understanding of energy balance. This knowledge will support the development of more precise nutritional interventions, environmentally adaptive breeding strategies, and individualized herd management practices, thereby enhancing both resilience and sustainability in dairy production.

4. References

Aernouts, B., E. Polshin, J. Lammertyn, and W. Saeys. 2011. Visible and near-infrared spectroscopic analysis of raw milk for cow health monitoring: Reflectance or transmittance? J. Dairy Sci. 94:5315–5329. https://doi.org/10.3168/jds.2011-4354.

Allen, M. S., B. J. Bradford, and M. Oba. 2009. The hepatic oxidation theory of the control of feed intake and its application to ruminants. J. Anim. Sci. 87:3317–34. https://doi.org/10.2527/jas.2009-1779.

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.

Brito, L. F., N. Bedere, F. Douhard, H. R. Oliveira, M. Arnal, F. Peñagaricano, A. P. Schinckel, C. F. Baes, and F. Miglior. 2021. Review: Genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world. Anim. 15 (Suppl. 1):100292. https://doi.org/10.1016/j.animal.2021.100292.

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.

Butler, W. R. 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. Livest. Prod. Sci. 83:211–218. https://doi.org/10.1016/S0301-6226(03)00112-X.

Butler, W. R., and R. D. Smith. 1989. Interrelationships between energy balance and postpartum reproductive function in dairy cattle. J. Dairy Sci. 72:767–783. https://doi.org/10.3168/jds.S0022-0302(89)79169-4.

Churakov, M., J. Karlsson, A. E. 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.

Connor, E. E. 2015. Invited review: Improving feed efficiency in dairy production: Challenges and possibilities. Anim. 9:395–408. https://doi.org/10.1017/S1751731114002997.

Czaplicki, S. 2022. Contribution to the assessment of animal welfare through unsupervised methods based on data collected within the Walloon milk recording system. Master thesis. Gembloux Agro-Bio Tech, University of Liège, Liège, Belgium. http://hdl.handle.net/2268.2/15949.

Drackley, J. K. 1999. Biology of dairy cows during the transition period: The final frontier? J. Dairy Sci. 82:2259–2273. https://doi.org/10.3168/jds.S0022-0302(99)75474-3.

Franceschini, S., C. Grelet, J. Leblois, N. Gengler, H. Soyeurt, and GplusE Consortium. 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.

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.

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, G. 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. Anim. 13:649–658. https://doi.org/10.1017/S1751731118001751.

Grelet, C., P. Dardenne, H. Soyeurt, J.A. Fernandez, A. Vanlierde, F. Stevens, N. Gengler, and F. Dehareng. 2021. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. Methods 186:97–111. https://doi.org/10.1016/j.ymeth.2020.07.012.

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.

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.

Ingvartsen, K. L., and J. B. Andersen. 2000. Integration of metabolism and intake regulation: A review focusing on periparturient animals. J. Dairy Sci. 83:1573–1597. https://doi.org/10.3168/jds.S0022-0302(00)75029-6.

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.

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.

Lucy, M. C. 2001. Reproductive loss in high-producing dairy cattle: Where will it end? J. Dairy Sci. 84:1277–1293. https://doi.org/10.3168/jds.S0022-0302(01)70158-0.

Manafiazar, G., L. Goonewardene, F. Miglior, D. H. Crews Jr., J. A. Basarab, E. Okine, and Z. Wang. 2016. Genetic and phenotypic correlations among feed efficiency, production, and selected conformation traits in dairy cows. Animal. 10:381-389. https://doi.org/10.1017/S1751731115002281.

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 13:3097. https://doi.org/10.3390/ani13193097.

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.

McArt, J. A. A., D. V. Nydam, and G. R. Oetzel. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. J. Dairy Sci. 95:5056–5066. https://doi.org/10.3168/jds.2012-5443.

McParland, S., and D.P. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J. Dairy Sci. 99:4056–4070. https://doi.org/10.3168/jds.2015-10051.

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.

Ospina, P. A., D. V. Nydam, T. Stokol, and T. R. Overton. 2010. Associations of elevated nonesterified fatty acids and β-hydroxybutyrate concentrations with early lactation reproductive performance and milk production in transition dairy cattle in the northeastern United States. J. Dairy Sci. 93:1596–1603. https://doi.org/10.3168/jds.2009-2852.

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 12: 1830. https://doi.org/10.3390/genes12111830.

Pinedo, P. J., A. De Vries, and D. W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. J. Dairy Sci. 93:2250-2261. https://doi.org/10.3168/jds.2009-2572.

Rhoads, M. L., R. P. Rhoads, M. J. VanBaale, R. J. Collier, S. R. Sanders, W. J. Weber, B. A. Crooker, and L. H. Baumgard. 2009. Effects of heat stress and plane of nutrition on lactating Holstein cows: I. Production, metabolism, and aspects of circulating somatotropin. J. Dairy Sci. 92:1986–1997. https://doi.org/10.3168/jds.2008-1641.

Roche, J. R., D. P. Berry, and E. S. Kolver. 2006. Holstein-Friesian strain and feed effects on milk production, body weight, and body condition score profiles in grazing dairy cows. J. Dairy Sci. 89:3532-43. https://doi.org/10.3168/jds.S0022-0302(06)72393-1.

Santos, J. E. P., E. J. DePeters, P. W. Jardon, and J. T. Huber. 2001. Effect of prepartum dietary protein level on performance of primigravid and multiparous

Holstein dairy cows. J. Dairy Sci. 84:213–224. https://doi.org/10.3168/jds.S0022-0302(01)74471-2.

Song, Y., Z. Wang, C. Zhao, Y. Bai, C. Xia, and C. Xu. 2021. Effect of negative energy balance on plasma metabolites, minerals, hormones, cytokines, and ovarian follicular growth rate in Holstein dairy cows. J. Vet. Res. 65:361–368. https://doi.org/10.2478/jvetres-2021-0035.

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. Gjuvsland, 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. GSE 52: 37. https://doi.org/10.1186/s12711-020-00556-4.

Wathes, D. C., Z. Cheng, N. Bourne, V. J. Taylor, M. P. Coffey, and S. Brotherstone. 2007. Differences between primiparous and multiparous dairy cows in the interrelationships between metabolic traits, milk yield and body condition score in the periparturient period. Domest. Anim. Endocrinol. 33:203–225. https://doi.org/10.1016/j.domaniend.2006.05.004.

Xu, W., J. Vervoort, E. Saccenti, B. Kemp, R.J. van Hoeij, and A.T.M. van Knegsel. 2020. Relationship between energy balance and metabolic profiles in plasma and milk of dairy cows in early lactation. J. Dairy Sci. 103:4795–4805. https://doi.org/10.3168/jds.2019-17777.

6

Chapter VI Appendix

1. Peer-reviewed scientific publications

Hu, H., S. Franceschini, P. Lemal, H. Atashi, C. Grelet, Y. Chen, K. Wijnrocx, H. Soyeurt, and N. Gengler. 2025. Comparing genetic architecture of MIR-predicted energy balance, a novel energy deficiency score and several biomarkers. J. Dairy Sci. (in press). https://doi.org/10.3168/jds.2025-26876.

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, H., H. Atashi, S. Franceschini, P. Lemal, C. Grelet, Y. Chen, K. Wijnrocx, H. Soyeurt, and N. Gengler. 2025. Using single-step genome-wide association analyses to compare predicted negative energy balance and a novel energy deficiency score in early-lactation Holstein cows. JDS Communications.6:792-796. https://doi.org/10.3168/jdsc.2025-0778.

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.

2. Contributions to international conferences

Hu, H., H. Atashi, C. Grelet, K. Wijnrocx, S. Franceschini, H. Soyeurt, and N. Gengler. 2024. Genetic analysis of predicted negative energy balance and its biomarkers of first-parity Holstein cows in early lactation. The 75th EAAP Annual Meeting. Florence, Italy. (poster)

Hu, H., Y. Chen, C. Grelet, and N. Gengler. 2023. Genetic parameters analysis of milk citrate for Holstein cows in early lactation. The 74th EAAP Annual Meeting. Lyon, France. (oral)

Hu, H., H. Atashi, Y. Chen, C. Grelet, S. Franceschini, H. Soyeurt, and N. Gengler. 2024. Comparing genetic architecture of predicted negative energy balance and its biomarkers of first-parity Holstein cows in early lactation. 1st Ph.D Symposium. Gembloux, Belgium. (poster)

7

Chapter VII Annexes

Table A-1. Pearson correlations between the LPNEB, LEDS, 15 biomarkers and 3 production traits (below diagonal, n=30,634), between the LRNEB, LPNEB, LEDS, 15 biomarkers and 3 production traits (above diagonal, n=965).

	LPNEB	LEDS	LB_BHB	NEFA	LIGF-1	GLU	LM_BHB	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
LRNEB	0.60	0.52	0.32	0.42	-0.24	-0.26	0.08	0.26	0.36	-0.11	-0.13	-0.03	0.41	0.48	0.11	-0.07	0.49	0.27	-0.12	0.01
LPNEB	0.00	0.75	0.47	0.64	-0.34	-0.39	0.09	0.36	0.54	-0.20	-0.26	-0.12	0.59	0.70	0.12	-0.18	0.71	0.33	-0.19	-0.07
LEDS	0.80	0.00	0.70	0.68	-0.34	-0.56	0.30	0.36	0.73	-0.47	-0.47	-0.26	0.44	0.69	-0.09	-0.33	0.65	0.19	-0.49	-0.08
LB_BHB	0.47	0.70	0.00	0.62	-0.35	-0.74	0.54	0.55	0.90	-0.32	-0.30	-0.12	0.38	0.51	0.03	-0.18	0.46	0.17	-0.49	-0.06
NEFA	0.79	0.78	0.58	0.00	-0.08	-0.59	0.24	0.41	0.69	-0.37	-0.52	-0.38	0.41	0.67	-0.06	-0.43	0.66	0.12	-0.05	0.11
LIGF-1	-0.58	-0.57	-0.48	-0.67	0.00	0.55	-0.39	-0.26	-0.45	0.36	0.04	-0.29	-0.39	-0.26	0.10	-0.15	-0.23	-0.26	0.33	0.52
GLU	-0.54	-0.67	-0.69	-0.72	0.86	0.00	-0.55	-0.26	-0.76	0.48	0.38	0.18	-0.30	-0.36	0.14	0.27	-0.35	-0.04	0.32	0.11
LM_BHB	0.36	0.60	0.70	0.61	-0.44	-0.59	0.00	0.24	0.54	-0.28	-0.02	0.25	0.24	0.20	0.03	0.14	0.18	0.23	-0.15	-0.23
CIT	0.36	0.32	0.42	0.53	-0.31	-0.33	0.57	0.00	0.62	-0.15	-0.21	-0.03	0.39	0.44	-0.02	-0.09	0.38	0.17	-0.16	-0.12
LACE	0.62	0.78	0.82	0.79	-0.61	-0.76	0.81	0.61	0.00	-0.50	-0.44	-0.18	0.47	0.57	-0.07	-0.29	0.54	0.15	-0.43	-0.12
C10:0	-0.44	-0.63	-0.48	-0.59	0.49	0.63	-0.52	-0.25	-0.60	0.00	0.88	0.58	0.17	-0.08	0.80	0.72	0.01	0.50	0.38	-0.02
C14:0	-0.36	-0.51	-0.42	-0.50	0.32	0.52	-0.34	-0.25	-0.48	0.94	0.00	0.87	0.28	-0.06	0.80	0.95	0.02	0.66	0.33	-0.24
C16:0	-0.06	-0.15	-0.21	-0.04	-0.04	0.21	0.11	0.09	-0.07	0.60	0.79	0.00	0.43	0.13	0.69	0.98	0.17	0.77	0.19	-0.41
C18:0	0.63	0.52	0.18	0.62	-0.42	-0.31	0.39	0.37	0.51	-0.07	0.11	0.47	0.00	0.83	0.54	0.38	0.90	0.84	0.09	-0.39
C18:1cis-9	0.76	0.79	0.39	0.83	-0.54	-0.53	0.50	0.40	0.67	-0.42	-0.26	0.17	0.82	0.00	0.28	0.09	0.98	0.67	0.05	-0.21
SCFA	-0.09	-0.32	-0.30	-0.22	0.19	0.35	-0.23	-0.06	-0.23	0.85	0.89	0.79	0.36	-0.03	0.00	0.75	0.40	0.78	0.31	-0.16
MCFA	-0.18	-0.28	-0.28	-0.22	0.11	0.34	-0.08	-0.05	-0.24	0.77	0.92	0.96	0.34	0.03	0.86	0.00	0.13	0.77	0.25	-0.33
LCFA	0.72	0.70	0.30	0.78	-0.47	-0.45	0.44	0.38	0.61	-0.28	-0.13	0.27	0.91	0.97	0.15	0.14	0.00	0.72	0.09	-0.22
FP	0.40	0.30	0.02	0.39	-0.25	-0.09	0.23	0.22	0.26	0.31	0.50	0.78	0.84	0.68	0.67	0.72	0.78	0.00	0.22	-0.38
PP	-0.17	-0.41	-0.58	-0.09	0.07	0.17	-0.29	0.01	-0.29	0.33	0.34	0.31	0.23	0.06	0.42	0.33	0.18	0.34	0.00	0.00
MY	0.20	0.06	0.14	0.05	0.10	0.00	-0.02	-0.02	0.02	-0.02	-0.07	-0.16	-0.06	0.01	-0.08	-0.12	-0.02	-0.09	-0.21	0.00

Abbreviations: LRNEB = logarithm measured probability negative energy balance; LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

Table A-2. The P-values of pearson correlations between the LPNEB, LEDS, 15 biomarkers and 3 production traits (below diagonal, n=30,634), between the LRNEB, LPNEB, LEDS, 15 biomarkers and 3 production traits (above diagonal, n=965).

	LPNEB	LEDS	LB_BHB	NEFA	LIGF-1	GLU	LM_BHB	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
LRNEB	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.78
LPNEB		0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
LEDS	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
LB_BHB	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.08
NEFA	0.00	0.00	0.00		0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.12	0.00
LIGF-1	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GLU	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00
LM_BHB	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.59	0.00	0.00	0.00	0.41	0.00	0.00	0.00	0.00	0.00
CIT	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.43	0.00	0.00	0.57	0.01	0.00	0.00	0.00	0.00
LACE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
C10:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.01	0.00	0.00	0.78	0.00	0.00	0.56
C14:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.07	0.00	0.00	0.59	0.00	0.00	0.00
C16:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C18:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.01	0.00
C18:1cis-9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.01	0.00	0.00	0.14	0.00
SCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00
MCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00
LCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
FP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
PP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.95
MY	0.00	0.00	0.00	0.00	0.00	0.52	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	

Abbreviations: LRNEB = logarithm measured probability negative energy balance; LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

Table A-3. The ranges of standard errors (SE) of heritability (diagonal), genetic correlations (above the diagonal), and phenotypic correlations (below the diagonal) among LRNEB, LPNEB, LEDS, 15 biomarkers, and 3 production traits are reported.

	LPNEB	LEDS	LB_BHB	NEFA	LIGF-1	GLU	LM_BHB	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
LPNEB	0.02	0.00	0.04	0.02	0.04	0.05	0.05	0.05	0.04	0.05	0.05	0.06	0.05	0.03	0.06	0.06	0.04	0.06	0.04	0.05
LEDS	0.02	0.02	0.03	0.02	0.04	0.04	0.04	0.05	0.03	0.04	0.05	0.06	0.06	0.02	0.06	0.06	0.04	0.06	0.04	0.06
LB_BHB	0.01	0.00	0.02	0.02	0.04	0.02	0.02	0.04	0.01	0.04	0.04	0.05	0.06	0.05	0.05	0.05	0.06	0.06	0.04	0.05
NEFA	0.00	0.00	0.00	0.02	0.03	0.03	0.03	0.04	0.02	0.04	0.05	0.06	0.05	0.03	0.05	0.06	0.04	0.06	0.04	0.05
LIGF-1	0.01	0.00	0.00	0.00	0.02	0.02	0.05	0.05	0.04	0.05	0.05	0.06	0.06	0.06	0.05	0.05	0.06	0.06	0.04	0.05
GLU	0.01	0.00	0.00	0.00	0.00	0.02	0.04	0.05	0.03	0.04	0.04	0.05	0.07	0.06	0.05	0.05	0.06	0.06	0.04	0.05
LM_BHB	0.01	0.00	0.00	0.00	0.01	0.00	0.02	0.03	0.02	0.05	0.05	0.06	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.06
CIT	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.03	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.04	0.05
LACE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.04	0.04	0.06	0.06	0.04	0.05	0.05	0.05	0.06	0.04	0.05
C10:0	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.02	0.01	0.03	0.06	0.06	0.01	0.02	0.06	0.04	0.04	0.05
C14:0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.02	0.05	0.06	0.01	0.01	0.06	0.03	0.04	0.05
C16:0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.02	0.04	0.06	0.02	0.00	0.06	0.02	0.05	0.05
C18:0	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.03	0.05	0.04	0.02	0.02	0.06	0.06
C18:1cis-9	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.00	0.04	0.06	0.06
SCFA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.06	0.02	0.01	0.06	0.02	0.04	0.05
MCFA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.06	0.00	0.02	0.06	0.02	0.04	0.05
LCFA	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.03	0.06	0.06
FP	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.06
PP	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.05
MY	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02

Abbreviations: LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

Table A-4. The P-values of heritability (diagonal), genetic correlations (above the diagonal), and phenotypic correlations (below the diagonal) among LRNEB, LPNEB, LEDS, 15 biomarkers, and 3 production traits are reported.

	LPNEB	LEDS	LB_BHB	NEFA	LIGF-1	GLU	LM_BHE	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
LPNEB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62	0.00	0.00	0.13	0.10	0.00	0.00	0.00	0.00
LEDS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
LB_BHB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.51	0.00	0.00
NEFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
LIGF-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.51	0.02	0.00	0.00	0.07	0.00	0.62	0.00	0.00
GLU	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.00	0.03	0.00	0.00
LM_BHB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.51	0.00	0.00	0.00	0.02	0.00	0.24	0.00	0.00
CIT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.05	0.07	0.32	0.07	0.00	0.62	0.02	0.00	0.16
LACE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.87	0.00	0.00
C10:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00
C14:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.10	0.00	0.00	0.00
C16:0	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
C18:0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.74
C18:1cis-9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
SCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.00	0.00	0.00	0.00	0.00	0.00
MCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.32	0.00	0.00	0.00	0.00	0.00	0.00
LCFA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.00
FP	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
PP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00
MY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Abbreviations: LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

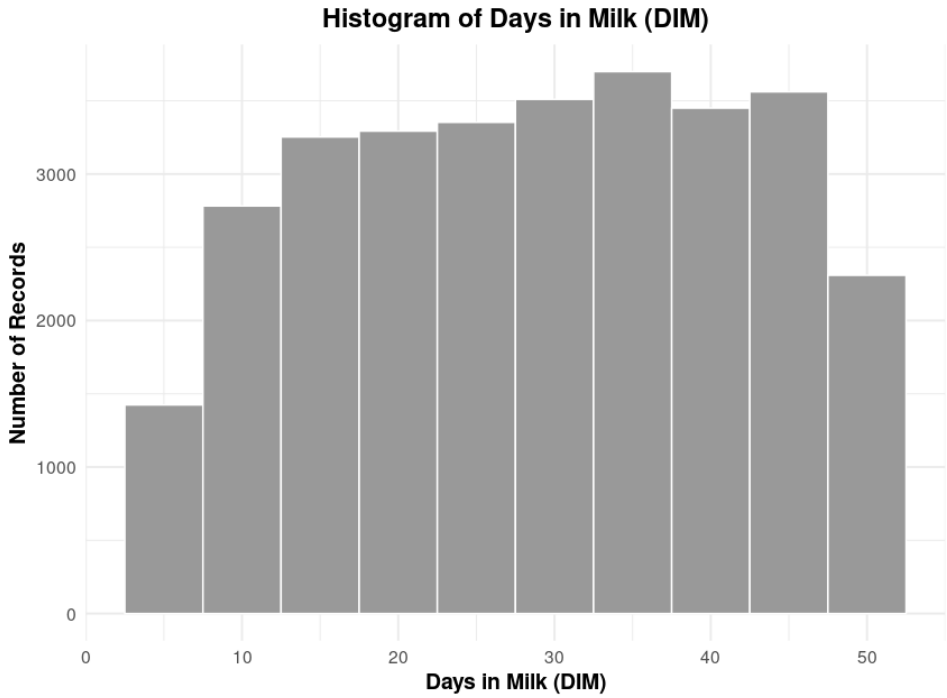


Figure A-1. Histogram of Days in Milk (DIM).Distribution of test-day records across the DIM range (5–50 d) used in this study.

Table A-5. Genomic correlations between LPNEB and the analyzed traits across individual Bos taurus chromosomes (BTA1–BTA29). Significant correlations were identified using false discovery rate (FDR) adjusted P-values (threshold: FDR < 0.05).

BTA	LEDS	LB_BHB	NEFA	LIGF-1	GLU	LM_BHB	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
BTA1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.95E-02	6.01E-02	1.00E+00	0.00E+00	0.00E+00	0.00E+00	7.55E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA2	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	7.15E-02	5.00E-04	0.00E+00	0.00E+00	0.00E+00	1.05E-01	0.00E+00	0.00E+00	0.00E+00	1.00E+00	2.00E-04
BTA3	0.00E+00	0.00E+00	0.00E+00	1.00E+00	4.35E-02	0.00E+00	1.00E+00	0.00E+00	4.00E-04	9.81E-02	1.19E-01	1.00E+00	0.00E+00	0.00E+00	1.16E-01	8.10E-03	0.00E+00	0.00E+00	0.00E+00
BTA4	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.50E-03	1.18E-01	0.00E+00	0.00E+00	1.00E+00	8.24E-02
BTA5	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.76E-02	7.52E-02	0.00E+00	1.00E+00	0.00E+00	0.00E+00	1.26E-01	0.00E+00	0.00E+00	1.24E-01	1.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00
BTA6	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA7	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.80E-03	0.00E+00	0.00E+00	1.00E+00	1.00E+00	0.00E+00	7.00E-04
BTA8	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.34E-01	0.00E+00	0.00E+00	0.00E+00	1.68E-01	0.00E+00	0.00E+00	0.00E+00	1.65E-01	2.30E-02	2.30E-03	8.68E-02	0.00E+00
BTA9	0.00E+00	3.33E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.10E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.10E-03	0.00E+00	0.00E+00	0.00E+00	1.09E-01	0.00E+00	0.00E+00
BTA10	0.00E+00	0.00E+00	0.00E+00	3.49E-02	0.00E+00	0.00E+00	1.14E-01	0.00E+00	1.10E-03	1.20E-01	0.00E+00	0.00E+00	0.00E+00	3.00E-03	2.80E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA11	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.92E-02	5.00E-04
BTA12	0.00E+00	1.00E-04	2.00E-04	4.49E-02	3.00E-04	1.17E-01	1.40E-01	3.77E-02	0.00E+00	1.00E+00	0.00E+00	6.35E-02	0.00E+00	1.70E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.70E-03
BTA13	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.50E-03	0.00E+00	1.00E-04	2.00E-04	0.00E+00	0.00E+00	2.36E-01	1.08E-01	0.00E+00	0.00E+00	3.00E-04	0.00E+00	7.90E-03	0.00E+00	0.00E+00
BTA14	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.89E-02	0.00E+00	0.00E+00	0.00E+00	7.00E-03	0.00E+00	0.00E+00
BTA15	0.00E+00	1.00E+00	1.00E+00	0.00E+00	1.00E+00	1.19E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	7.00E-04	0.00E+00
BTA16	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.61E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.20E-03	0.00E+00
BTA17	1.00E+00	0.00E+00	2.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.32E-02	0.00E+00	0.00E+00	1.00E+00	4.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.61E-01	0.00E+00	0.00E+00
BTA18	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.30E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA19	0.00E+00	5.00E-04	0.00E+00	4.00E-04	1.29E-01	5.40E-03	0.00E+00	4.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA20	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.26E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA21	2.00E-04	0.00E+00	0.00E+00	0.00E+00	1.13E-01	0.00E+00	0.00E+00	0.00E+00	1.72E-01	8.20E-03	1.00E+00	1.00E+00	1.00E+00	0.00E+00	1.00E+00	4.53E-02	1.00E+00	0.00E+00	1.00E+00
BTA22	0.00E+00	0.00E+00	0.00E+00	5.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.47E-01	1.30E-03	0.00E+00	2.46E-01	1.87E-02	0.00E+00	0.00E+00	0.00E+00	1.95E-01
BTA23	0.00E+00	0.00E+00	1.00E+00	1.00E+00	5.50E-03	0.00E+00	0.00E+00	0.00E+00	3.00E-04	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA24	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-04	2.00E-04	0.00E+00	0.00E+00	2.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.55E-01	0.00E+00
BTA25	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.79E-01	0.00E+00	0.00E+00	0.00E+00	1.03E-01	0.00E+00	0.00E+00	0.00E+00
BTA26	2.10E-03	0.00E+00	0.00E+00	1.00E+00	0.00E+00	1.00E-04	0.00E+00	0.00E+00	2.00E-04	2.52E-01	1.10E-03	0.00E+00	0.00E+00	1.10E-03	1.00E+00	0.00E+00	0.00E+00	0.00E+00	2.33E-01
BTA27	1.00E+00	0.00E+00	0.00E+00	1.40E-01	0.00E+00	2.43E-01	0.00E+00	1.27E-01	0.00E+00	0.00E+00	0.00E+00	1.77E-01	0.00E+00	0.00E+00	0.00E+00	5.00E-04	0.00E+00	0.00E+00	0.00E+00
BTA28	1.00E+00	0.00E+00	4.00E-03	0.00E+00	1.90E-01	0.00E+00	0.00E+00	0.00E+00	6.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.04E-02	0.00E+00
BTA29	0.00E+00	1.37E-01	0.00E+00	2.40E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00

Abbreviations: BTA = Bos taurus chromosome; FDR=false discovery rate; LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

Table A-6. Mean absolute genomic correlation (MeanAbsCorr) values for LPNEB calculated for each Bos taurus autosome.

Bos taurus autosomes	MeanAbsCorr	All_BTA	P_value	FDR
BTA1	0.38	0.43	0.00E+00	0.00E+00
BTA2	0.44	0.43	8.94E-03	1.30E-02
BTA3	0.43	0.43	4.05E-01	4.19E-01
BTA4	0.46	0.43	5.04E-10	1.22E-09
BTA5	0.43	0.43	2.61E-01	2.91E-01
BTA6	0.45	0.43	1.23E-05	2.09E-05
BTA7	0.48	0.43	0.00E+00	0.00E+00
BTA8	0.39	0.43	3.00E-10	7.90E-10
BTA9	0.46	0.43	1.23E-11	4.48E-11
BTA10	0.39	0.43	5.14E-11	1.49E-10
BTA11	0.44	0.43	4.28E-02	5.91E-02
BTA12	0.43	0.43	2.36E-01	2.74E-01
BTA13	0.44	0.43	8.05E-02	1.02E-01
BTA14	0.35	0.43	0.00E+00	0.00E+00
BTA15	0.42	0.43	1.31E-01	1.58E-01
BTA16	0.39	0.43	2.63E-07	5.86E-07
BTA17	0.47	0.43	2.78E-11	8.97E-11
BTA18	0.48	0.43	0.00E+00	0.00E+00
BTA19	0.45	0.43	8.42E-06	1.53E-05
BTA20	0.48	0.43	0.00E+00	0.00E+00
BTA21	0.42	0.43	4.19E-01	4.19E-01
BTA22	0.48	0.43	2.22E-16	9.20E-16
BTA23	0.39	0.43	5.00E-07	1.04E-06
BTA24	0.41	0.43	7.04E-02	9.28E-02
BTA25	0.51	0.43	0.00E+00	0.00E+00
BTA26	0.40	0.43	1.52E-04	2.32E-04
BTA27	0.39	0.43	3.48E-05	5.60E-05
BTA28	0.43	0.43	2.85E-01	3.06E-01
BTA29	0.46	0.43	9.96E-07	1.92E-06

Abbreviations: BTA = Bos taurus autosome; MeanAbsCorr = Mean absolute genomic correlation; All_BTA = Combined results across all Bos taurus autosomes; FDR= False discovery rate; LPNEB = Logarithm probability negative energy balance predicted by mid-infrared spectra.

Table A-7. Genomic correlations between LEDS and the analyzed traits across individual Bos taurus chromosomes (BTA1–BTA29). Significant correlations were identified using false discovery rate (FDR) adjusted P-values (threshold: FDR < 0.05).

BTA	LPNEB	LB_BHB	NEFA	LIGF-1	GLU	LM_BHB	CIT	LACE	C10:0	C14:0	C16:0	C18:0	C18:1cis-9	SCFA	MCFA	LCFA	FP	PP	MY
BTA1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-02	0.00E+00	7.00E-02	0.00E+00	0.00E+00	6.00E-02	2.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA2	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	9.00E-02
BTA3	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.00E-02	8.00E-04	0.00E+00	0.00E+00	6.00E-02	0.00E+00	2.00E-04	1.00E+00	0.00E+00	9.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA4	0.00E+00	9.00E-04	0.00E+00	0.00E+00	0.00E+00	2.00E-02	8.00E-04	1.00E+00	0.00E+00	5.00E-02	0.00E+00	0.00E+00	0.00E+00	7.00E-04	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00
BTA5	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.00E-02	0.00E+00	3.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	7.00E-04	1.10E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA6	0.00E+00	7.00E-04	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA7	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.00E-04	3.00E-02	1.00E+00	1.40E-01	0.00E+00	0.00E+00	0.00E+00	1.20E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA8	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	7.00E-04	0.00E+00	0.00E+00	1.00E+00	1.40E-01	0.00E+00	1.00E+00	0.00E+00	1.00E+00	1.00E+00
BTA9	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.20E-01	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.30E-01
BTA10	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	1.20E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA11	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-04	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	1.10E-01
BTA12	0.00E+00	0.00E+00	1.00E-04	0.00E+00	0.00E+00	1.00E-04	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA13	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	2.30E-01	0.00E+00	0.00E+00	0.00E+00	1.00E-02	4.00E-02	1.00E-02	0.00E+00	0.00E+00
BTA14	0.00E+00	0.00E+00	0.00E+00	2.00E-04	0.00E+00	8.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-02	0.00E+00	0.00E+00	1.00E+00	1.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA15	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	1.00E+00	0.00E+00	9.00E-04	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA16	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	6.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	7.00E-02	1.00E-02
BTA17	1.00E+00	1.00E-02	0.00E+00	5.00E-02	0.00E+00	7.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	7.00E-02	0.00E+00
BTA18	0.00E+00	0.00E+00	3.00E-04	2.00E-04	0.00E+00	3.00E-04	0.00E+00	0.00E+00	5.00E-04	0.00E+00	2.00E-04	0.00E+00	1.00E+00	1.00E+00	1.00E-02	0.00E+00	1.00E-02	0.00E+00	0.00E+00
BTA19	0.00E+00	0.00E+00	4.00E-04	2.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	9.00E-02	2.00E-04
BTA20	0.00E+00	0.00E+00	3.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.00E-04	1.30E-01	1.00E-02	1.00E+00	0.00E+00	1.80E-01	1.00E-02	3.00E-02	1.00E-04	1.00E+00	1.00E-04
BTA21	2.00E-04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	1.30E-01	2.20E-01	1.40E-01	0.00E+00	0.00E+00	2.00E-01	0.00E+00	2.10E-01	8.00E-02	0.00E+00
BTA22	0.00E+00	0.00E+00	0.00E+00	8.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.50E-01	1.60E-01	0.00E+00	2.10E-01	1.00E-02	0.00E+00	2.00E-04	0.00E+00	0.00E+00
BTA23	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA24	1.00E+00	2.00E-02	1.00E+00	0.00E+00	0.00E+00	6.00E-04	2.30E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA25	0.00E+00	0.00E+00	0.00E+00	1.40E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.20E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.00E-04	1.00E+00	1.60E-01	0.00E+00
BTA26	0.00E+00	0.00E+00	0.00E+00	9.00E-04	0.00E+00	1.00E+00	5.00E-04	0.00E+00	1.20E-01	0.00E+00	0.00E+00	2.00E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA27	1.00E+00	0.00E+00	0.00E+00	0.00E+00	9.00E-02	0.00E+00	0.00E+00	1.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00
BTA28	1.00E+00	0.00E+00	0.00E+00	0.00E+00	9.00E-02	0.00E+00	0.00E+00	0.00E+00	1.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
BTA29	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.00E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.80E-01

Abbreviations: BTA = Bos taurus chromosome; FDR = false discovery rate; LPNEB = logarithm probability negative energy balance predicted by mid-infrared spectra; LEDS = logarithm probability energy deficit score; LB_BHB =log10 blood beta-hydroxybutyrate acid predicted by mid-infrared spectra; NEFA= blood non-esterified fatty acids predicted by mid-infrared spectra; LIGF-1= log10 blood insulin-like growth factor 1 predicted by mid-infrared spectra; GLU = blood glucose predicted by mid-infrared spectra; LM_BHB = log10 milk beta-hydroxybutyrate acid predicted by mid-infrared spectra; CIT = milk citrate predicted by mid-infrared spectra; LACE = log10 milk acetone predicted by mid-infrared spectra; C10:0 = milk decanoic acid predicted by mid-infrared spectra; C14:0 = milk myristic acid predicted by mid-infrared spectra; C16:0 = milk palmitic acid predicted by mid-infrared spectra; C18:0 = milk stearic acid predicted by mid-infrared spectra; C18:1 cis-9 = milk oleic acid predicted by mid-infrared spectra; SCFA = milk short-chain fatty acids predicted by mid-infrared spectra; MCFA = milk medium chain fatty acids predicted by mid-infrared spectra; LCFA = milk long-chain fatty acids predicted by mid-infrared spectra; FP = milk fat percentage predicted by mid-infrared spectra; PP = milk protein percentage predicted by mid-infrared spectra; MY = milk yield.

Table A-8. Mean absolute genomic correlation (MeanAbsCorr) values for LEDS calculated for each Bos taurus autosome.

Bos taurus autosomes	MeanAbsCorr	All_BTA	P_value	FDR
BTA1	0.48	0.50	1.25E-08	4.04E-08
BTA2	0.53	0.50	2.55E-14	1.48E-13
BTA3	0.50	0.50	3.09E-01	3.44E-01
BTA4	0.51	0.50	1.11E-01	1.29E-01
BTA5	0.50	0.50	8.24E-01	8.24E-01
BTA6	0.51	0.50	2.04E-02	2.81E-02
BTA7	0.52	0.50	4.62E-07	1.22E-06
BTA8	0.50	0.50	3.69E-01	3.97E-01
BTA9	0.53	0.50	2.10E-09	7.63E-09
BTA10	0.49	0.50	3.55E-02	4.68E-02
BTA11	0.48	0.50	8.67E-06	1.68E-05
BTA12	0.48	0.50	1.22E-03	1.86E-03
BTA13	0.53	0.50	7.45E-06	1.54E-05
BTA14	0.43	0.50	0.00E+00	0.00E+00
BTA15	0.47	0.50	5.66E-10	2.35E-09
BTA16	0.53	0.50	1.63E-08	4.74E-08
BTA17	0.52	0.50	1.06E-03	1.71E-03
BTA18	0.53	0.50	1.42E-06	3.16E-06
BTA19	0.58	0.50	0.00E+00	0.00E+00
BTA20	0.49	0.50	1.06E-01	1.28E-01
BTA21	0.50	0.50	7.23E-01	7.49E-01
BTA22	0.54	0.50	1.67E-13	8.07E-13
BTA23	0.47	0.50	7.78E-07	1.88E-06
BTA24	0.47	0.50	9.09E-05	1.55E-04
BTA25	0.57	0.50	0.00E+00	0.00E+00
BTA26	0.47	0.50	1.95E-05	3.54E-05
BTA27	0.43	0.50	0.00E+00	0.00E+00
BTA28	0.52	0.50	6.44E-03	9.34E-03
BTA29	0.51	0.50	5.55E-02	6.99E-02

Abbreviations: BTA = Bos taurus autosome; MeanAbsCorr = Mean absolute genomic correlation; All_BTA = Combined results across all Bos taurus autosomes; FDR= False discovery rate; LEDS = Logarithm probability energy deficit score.