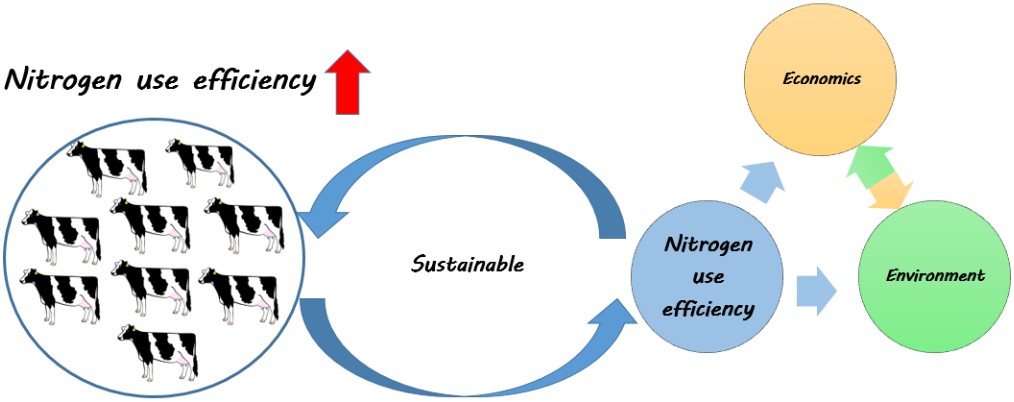
LIÈGE université
Gembloux
Agro-Bio Tech

Genetic analyses of different nitrogen use efficiency proxies and their relationships with other traits for Holstein cows

Yansen Chen



COMMUNAUTÉ FRANÇAISE DE BELGIQUE

UNIVERSITÉ DE LIÈGE – GEMBLOUX AGRO-BIO TECH

Genetic analyses of different nitrogen use efficiency proxies and their relationships with other traits for Holstein cows

Yansen CHEN

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

Promoteurs: Prof. Nicolas GENGLER & Dr. Sylvie VANDERICK

Année civile: 2023

Abstract

The definition of nitrogen (N) use efficiency (NUE) for lactating dairy cows traditionally is milk N divided by N intake. The NUE of dairy cows is associated negatively with N pollution and positively with economic results of dairy producers. The N pollution from dairy cows refers to N from manure and urine, which are mainly produced from undigested protein from the feed. Protein is the most expensive component of dairy cattle feed, and its loss reduces the profits of dairy farms. Therefore, measuring and improving the NUE of dairy cows can promote the sustainable economic development of dairy production and strengthen its social acceptability. Although the NUE in dairy cows can be affected by many factors, the economic importance of genetic improvement for efficiency traits in cattle is recognized by dairy producers.

The ultimate objective of this study was to contribute to the breeding of dairy cows that are both more economical and more respectful of the environment. However, the NUE is difficult to measure, which is why few studies have been conducted on genetic selection for NUE in dairy cows. So, the use of proxies of NUE to conduct genetic selection could be a good choice for its ease to be measured on a large scale. Different proxies of NUE were genetically analyzed in this thesis and their genetic relationships with other already selected traits were explored in Holstein cows in the Walloon Region of Belgium.

In the first step, milk urea concentration (MU), as the traditional proxy of NUE, was genetically analyzed, and its genetic correlations with 11 other traits of economic interest were estimated. The results showed that the average daily heritability and repeatability of MU in the first 3 parities ranged from 0.19 to 0.22 and from 0.47 to 0.48, respectively. High genetic correlations (0.94 to 0.97) were found among MU in the first 3 parities. The genetic correlation between MU and the 11 traits of interest ranged from −0.28 (milk yield) to 0.28 (somatic cell score).

Then, predict NUE (PNUE) and N losses (PNL), as alternative and novel proxies for NUE, predicted by milk mid-infrared spectra, milk yield, and parity. Genetic analyses were performed on these two phenotypes and genetic correlations were estimated between them and 30 other traits of interest. The estimated heritability and repeatability of PNUE and PNL in primiparous and multiparous ranged from 0.12 to 0.14, and from 0.40 to 0.55, respectively. The approximate genetic correlations between PNUE and 30 traits of interest ranged from −0.46 (udder depth) to 0.47 (milk yield). Additionally, the approximate genetic correlations between PNL, lower values representing less N pollution, and 30 traits of interest ranged from −0.32 (angularity) to 0.57 (direct calving ease).

Since the genetic correlations estimated between MU and PNUE were very low; the estimated breeding values (EBV) of the three features (N intake (NINT), milk true protein N (MTPN), and milk urea N yield (MUNY)) were used to build a new N efficiency index (NEI), the purpose of which was to quantitatively combine MU and PNUE. The approximate genetic correlations between the NEI and 37 other traits (economic indices) of interest were investigated. The NEI showed positive genetic correlations with production yield traits (0.08 to 0.46), and negative genetic correlations with the investigated functional traits and economic indices (−0.71 to −0.07), except for production and functional type economic indices. In addition, increasing NEI in early lactation favors a reduction in the intensity of methane emissions and increases dry matter intake but is detrimental to energy balance (given in general discussion). We then explored the potential impact of genetic selection for NEI on bulls. The daughters of the bulls with higher NEI had lower NINT and MUNY, and higher MTPN.

Genomic selection is commonly applied in animal breeding, so whether NEI and its composition traits can be used for genomic prediction was verified to select at earlier stage the dairy cows and bulls. The prediction accuracies of the NEI and its composition traits performed, using single-step genomic best linear unbiased prediction (ssGBLUP) analyses, varied from 0.48 to 0.66 for genotyped cows, from 0.35 to 0.55 for non-genotyped cows, and from 0.39 to 0.56 for bulls.

Finally, we investigated the genomic background of NEI to understand better its genetic variability. So, the genomic regions associated with NEI and its composition traits were identified and the functional annotation of the identified genomic regions was analyzed. The largest explanatory genomic region of NEI was located at position 1.52-2.09 Mb of Bos taurus autosome 14, which explained 0.58% of the total additive genetic variance. The 16 key candidate genes were identified as related to NEI and its composition traits, which are mainly expressed in the milk cells, mammary, and liver tissues. Annotated quantitative trait loci (QTLs) are mostly reported to be related to milk, health, and production traits based on the identified genomic regions.

In conclusion, this study showed that it is possible to develop genetic selection for dairy cows that are both more economical and more respectful of the environment. Moreover, the developed NEI has the advantage of available phenotypes through large-scale prediction. Therefore, the NEI has the potential for routine application in dairy cattle breeding in the future, contributing a novel possibility to reduce N pollution and improve economic results for dairy producers.

Résumé

L’efficience de l’utilisation de l'azote (NUE) pour les vaches lactantes est généralement définie comme étant l’azote du lait divisé par l'azote ingérée. La NUE des vaches laitières est associée négativement à la pollution azotée et positivement aux résultats économiques des producteurs laitiers. La pollution azotée des vaches laitières est due à l’azote contenue dans le fumier et l’urine, azote provenant principalement des protéines alimentaires non digérées. Les protéines étant l’ingrédient le plus coûteux de l’alimentation des vaches laitières, leur perte diminue par conséquent les bénéfices des exploitations laitières. Dès lors, mesurer et améliorer le NUE des vaches laitières permettrait de promouvoir un développement économique durable la production laitière tout en renforçant son acceptabilité sociale. Malgré que de nombreux facteurs affectent la NUE des vaches laitières, les producteurs laitiers ont bien conscience de l’importance économique que constitue l’amélioration génétique des caractères d’efficience.

L'objectif final de cette étude a été de contribuer à la sélection de bovins laitiers à la fois plus économiques et respectueux de l'environnement. Cependant, vu la difficulté de mesurer le NUE, peu d’études concernant la sélection génétique de ce caractère ont été menées. Par conséquent, l'utilisation d'indicateurs de la NUE plus faciles à mesurer à grande échelle pourrait constituer un bon choix pour appliquer une sélection génétique. Des analyses génétiques de différents indicateurs de la NUE ont été effectuées et leurs relations génétiques avec d’autres caractères d’intérêt ont été étudiées parmi vaches Holstein wallonnes.

Dans un premier temps, la concentration en urée du lait (MU), qui est l’indicateur traditionnel de la NUE, a été analysée génétiquement, et ses corrélations génétiques avec 11 autres caractères d'intérêt économique ont été estimées. Les résultats ont montré que l'héritabilité et la répétabilité journalières moyennes de MU dans les 3 premières lactations varient de 0,19 à 0,22 et de 0,47 à 0,48, respectivement. Des corrélations génétiques élevées (0,94 à 0,97) ont été trouvées entre MU dans les 3 premières lactations. Les corrélations génétiques entre MU et les 11 caractères d'intérêt ont variées de −0,28 (production de lait) à 0,28 (score cellulaire).

Ensuite, en fonction du spectre moyen infrarouge du lait, du rendement laitier et de la parité, des prédictions de la NUE (PNUE) et des pertes d'azote (PNL) ont été réalisées afin de produire des indicateurs alternatifs de la NUE. Des analyses génétiques ont été effectuées sur ces deux indicateurs alternatifs et des corrélations génétiques ont été estimées entre eux et 30 autres caractères d'intérêt. Les héritabilités et les répétabilités estimées du PNUE et du PNL chez les primipares et les multipares se sont situées respectivement entre 0,12 et 0,14, et entre 0,40 et 0,55. Les corrélations génétiques approximatives entre la PNUE et 30 caractères d'intérêt ont variées de −0,46 (profondeur du pis) à 0,47 (production de lait). En outre, les corrélations génétiques approximatives entre PNL - les valeurs les plus basses représentant une

moindre pollution par l'azote - et les 30 caractères d'intérêt ont variées de −0,32 (angularité) à 0,57 (facilité de vêlage directe).

Les corrélations génétiques approximatives entre le PNUE et 30 traits d'intérêt ont variées de −0,46 (profondeur du pis) à 0,47 (rendement laitier). De plus, les corrélations génétiques approximatives entre PNL, les valeurs les plus basses représentant une moindre pollution par l'azote, et 30 traits d'intérêt ont variées de −0,32 (angularité) à 0,57 (facilité de vêlage direct).

Au vu des très faibles corrélations génétiques estimées entre MU et PNUE, un nouvel index d'efficience azotée (NEI), permettant de combiner quantitativement MU et PNUE, a été construit en combinant les valeurs d'élevage estimées de l'azote ingérée (NINT), de l'azote protéique du lait (MTPN) et de l'azote de l'urée du lait (MUNY). Les corrélations génétiques approximatives entre le NEI et 37 autres caractères d'intérêt (dont des index économiques) ont été calculées. Des corrélations génétiques positives avec les caractères de production (0,08 à 0,46), et des corrélations génétiques négatives avec les caractères fonctionnels et les index économiques étudiés (−0,71 à −0,07) ont été trouvées, excepté pour les index économiques de production laitière (V€L) et de type fonctionnel (V€T). De plus, le NEI favoriserait l'intensité de la réduction des émissions de méthane ainsi que l'augmentation de matière sèche ingérée en début de lactation mais au détriment de la balance énergétique. L'impact potentiel de la sélection génétique basée sur le NEI sur les taureaux a été investigué et a montré que les filles des taureaux présentant un NEI élevé possèdent des NINT et des MUNY plus faibles, et des MTPN plus élevées.

La sélection génomique étant devenue une pratique courante au sein des animaux d'élevage, les prédictions génomiques du NEI ainsi que de ses composants ont été testées à des fins de sélection précoce. Les précisions de ces prédictions génomiques en utilisant la méthode « single step genomic BLUP » (ssGBLUP) varient de 0,48 à 0,66 pour les vaches génotypées, de 0,35 à 0,55 pour les vaches non génotypées et de 0,39 à 0,56 pour les taureaux.

Finalement, le contexte génomique du NEI a été étudié afin de mieux comprendre sa variabilité génétique. Ainsi, les régions génomiques associées au NEI et à ses composants ont été identifiées et l'annotation fonctionnelle des régions génomiques identifiées a été analysée. La plus grande région génomique explicative du NEI est située à la position 1,52-2,09 Mb de l'autosome 14 de Bos taurus et explique 0,58% de la variance génétique additive totale. Les 16 gènes candidats clés identifiés comme étant liés au NEI et à ses composants sont principalement exprimés dans les cellules du lait, les tissus mammaires et le foie. La plupart des loci quantitatifs (QTL) annotés ont été rapportés comme étant liés aux caractères de production du lait et de la santé d'après les régions génomiques identifiées.

En conclusion, cette étude a montré qu'il est possible de développer une sélection génétique pour des bovins laitiers à la fois plus économiques et respectueux de l'environnement. En outre, le NEI développé offre l'avantage de phénotypes disponibles grâce à la prédiction à grande échelle. Par conséquent, le NEI a le potentiel pour être appliqué en routine à l’avenir en bovins laitiers, contribuant de la sorte à une nouvelle possibilité de réduire la pollution azotée et d'améliorer les résultats économiques des producteurs laitiers.

Acknowledgments

Firstly, I would like to express my sincere gratitude to the China Scholarship Council (CSC) for supporting my living cost during my PhD. Without their financial support, I would not have been able to live in Belgium and complete my PhD project.

I am deeply grateful to my supervisors, Nicolas Gengler and Sylvie Vanderick. Although we did not work in the office for a year and a half during the Covid-19 period, we kept in close contact through regular teleconferences. Their continued guidance and assistance have made my work easy. Nicolas Gengler brought a lot of novel ideas to this thesis; Sylvie Vanderick taught me how to work collaboratively with other colleagues. They selflessly shared their knowledge to help with the completion of this thesis.

I would like to thank the members of my thesis committee: Hélène Soyeurt, Clément Grelet, Rodrigo R. Mota, Jérôme Bindelle, Martine Schroyen, and Sven König, for their insightful comments and suggestions. In particular, Clément provided me with good raw data, which facilitated the start and completion of this thesis.

Many thanks to Hadi for revising my article and discussing the technical details. Moreover, I gratefully acknowledge the co-authors of my publications for their help. I am also grateful to all the programs that funded my published papers.

I am also very grateful to all the members of our team: Alice, Hélène, Darlene, José, Katrien, Malak, and Pauline. I thank them for giving me a pleasant working atmosphere. Many thanks to my friend Axell and her family for their help during the time I lived in Belgium for more than three years. I remember that my friend, Hedi, was a very good guide for me to learn quantitative genetics.

我非常感谢在让布卢认识的所有朋友, 是你们让我博士生活更加丰富多彩。从最初以为让布卢只有很少国内学生, 独自生活半年; 到后来逐渐认识更多朋友, 回复“正常”生活; 接着新冠到来, 又开始了独自生活; 最后, 又在学联服务一年后, 圆满完成了自己的博士路程。需要感谢的人很多: 张玉萍, 张磊, 黄亚涛, 马玉玲, 赵雪梅, 王术, 丁宁, 肖世泽, 李永峰, 王爽, 李诺男, 蒋欣…

最后, 我需要特别感谢我的父母, 姐姐和妹妹。他们理解我没能三年回家, 并且鼓励我去完成自己的梦想。当然, 我最需要感谢我的女朋友—鸿晴。没有她的陪伴和支持, 我也不可能完成本论文。

06/01/2023, Gembloux

Table of contents

List of figures

List of tables

List of abbreviations

AGC	Approximate genetic correlation
ANG	Angularity
BCS	Bod condition score
BDE	Body depth
BLUP	Best linear unbiased prediction
BTA	Bos Taurus autosome
CWI	Chest width
DCE	Direct calving ease
DIM	Days in milk
DMI	Dry matter intake
EBV	Estimated breeding value
FAN	Foot angle
FE	Feed efficiency
FF	Female fertility
FP	Fat percentage
FTP	Front teat placement
FUD	Fore udder
FY	Fat yield
GEBV	Genomic estimated breeding value
GWAS	Genome-wide association study
h^2	Heritability
LMI	Log-transformed ratio of PME divided by the daily milk yield (kg/d) recorded
LONG	Longevity
MCE	Maternal calving ease
MIR	Mid-infrared
MTPN	Milk true protein nitrogen
MTPN1	Milk true protein nitrogen in primiparous cows
MTPN2+	Milk true protein nitrogen in multiparous cows
MU	Milk urea concentration
MUN	Milk urea nitrogen
MUNY	Milk urea nitrogen yield
MUNY1	Milk urea nitrogen yield in primiparous cows
MUNY2+	Milk urea nitrogen yield in multiparous cows
MY	Milk yield
N	Nitrogen

NEB	Negative energy balance
NEI	Nitrogen efficiency index
NINT	Nitrogen intake
NINT1	Nitrogen intake in primiparous cows
NINT2+	Nitrogen intake in multiparous cows
NL	Nitrogen losses
NPN	Non-protein-N except urea
NUE	Nitrogen use efficiency
OCS	Overall conformation score
OFL	Overall feet and leg score
OUS	Overall udder score
PNL	Predicted nitrogen losses
PNL1	Predicted nitrogen losses in first-parity Holstein cows
PNL2+	Predicted nitrogen losses in second- to fifth-parity Holstein cows
PNUE	Predicted nitrogen use efficiency
PNUE1	Predicted nitrogen use efficiency in first-parity Holstein cows
PNUE2+	Predicted nitrogen use efficiency in second- to fifth-parity Holstein cows
PME	Daily methane production using existing milk MIR equation
PP	Protein percentage
PROD	Production
PY	Protein yield
QTL	Quantitative trait locus
RAN	Rump angle
RLR	Rear leg rear view
RLS	Rear leg set
RTP	Rear teat placement
RUH	Rear udder height
RWI	Rump width
SCS	Somatic cell score
SD	Standard deviation
SE	Standard error
SNP	Single nucleotide polymorphism
ssGBLUP	Single-step genomic best linear unbiased prediction
ssGWAS	Single-step genome-wide association study
STA	Stature
TLE	Teat length
UDE	Udder depth
UH	Udder health

USU	Udder support
V€C	Capacity economic index
V€F	Functional economic index
V€G	Global economic index
V€L	Production economic index
V€M	Member economic index
V€P	Udder economic index
V€T	Functional type economic index

1

Chapter I General introduction

1. Introduction to high-efficiency dairy cows

Over the past few decades, the current dairy breeding goals have shifted from increasing milk production to more balanced breeding and have focused more on health, welfare, and environmental sustainability (Miglior et al., 2017). Among these traits, the environmental sustainability status generally refers to increased efficiency and reduced emissions. The definition of a high-efficiency dairy cow is measured by different metrics in different fields (e.g., economics, nutrition, and genetics) (de Ondarza and Tricarico, 2017). High-efficiency lactating Holstein cows in this study refer to cows with a high ratio of cow products to feed intake. Although efficiency traits in dairy cattle are affected by a variety of factors, the economic importance of genetic improvement for efficiency traits in cattle is recognized by the world's cattle producers (Brito et al., 2020). The used indicators to assess feed efficiency (FE) are currently dry matter intake (DMI) and residual feed intake (RFI). Brito et al. (2020) made a good review of the recent genetic research on these FE indicators. The U.S. genomic evaluation incorporated a new indicator, feed saved, to improve cow efficiency (Parker Gaddis et al., 2021). The recently completed international project "Efficient Dairy Genome Project" (https://genomedairy.ualberta.ca/) also promoted faster addition of FE to current dairy cattle breeding. However, the FE as a real phenotype is still difficult and expensive to measure, and researchers often use predictive methods or proxies to study it.

Potential future applications of high-throughput predictors for FE prediction in breeding programs are summarized by Siberski-Cooper and Koltes (2022). The authors focus on the application of sensor (leg, ear, rumen), imaging-based, and mid-infrared (MIR) spectral techniques for FE traits prediction. Cavani et al. (2022) reported that feeding behavior may be a good proxy for FE. Easier to-measure traits (milk composition, live weight, and milk MIR) for FE prediction and genetic selection in developing countries were reported by Madilindi et al. (2022). The FE traits of lactating dairy cows are usually only considered in the calculation process of DMI, milk production, and energy maintenance. However, the FE can be further divided into energy and nitrogen (N) efficiency. The N use efficiency (NUE) of lactating dairy cows is also an important economic and environmental factor but is often overlooked, especially in genetic selection.

The NUE is directly related to feed protein use and N pollution from dairy cows. Protein is the most expensive component of dairy cow's feed, and its loss reduces farmers profits. The N pollution from dairy cows refers to N from manure and urine, which enters groundwater as nitrate and is released to the atmosphere as nitrous oxide (N_2O) (Rochette et al., 2014). Livestock production contributes to 18% of the global human-induced greenhouse gases emissions in the form of carbon dioxide (CO_2), methane (CH_4), and N_2O (Moran and Wall, 2011). The N emissions from the dairy industry have been identified as an important factor causing groundwater and surface water pollution, as well as greenhouse gases emissions (Adenuga et al., 2019). Reducing no-CO_2 (CH_4 and N_2O) would be more beneficial to decrease the rate of climate change under conditions of limiting temperature increases to 2°C by 2100 (Reisinger and Clark, 2017). The EU has developed a strategy to reduce greenhouse gases emissions to net-zero by 2050, which also includes specific regulations on N_2O emissions from the agriculture industry (aan den Toorn et al., 2021). The main sources of N_2O production in cattle are deposited urine and manure (Rochette et al., 2014), which are mainly due to the unused protein in the food. Therefore, measuring and improving NUE is not only beneficial for the protection of the environment, but also beneficial to the farms. By doing so, we may promote the sustainable economic development of dairy production, and strengthen its social acceptability.

Spanghero and Kowalski (2021), based on the 86 N balance experiments, found that the average NUE in cattle is 27% suggesting that the NUE in dairy cattle may have great potential for improvement. The NUE, like FE traits, is influenced by many factors such as feed and management (Jonker et al., 2002; Kidane et al., 2018; Herremans et al., 2020). The NUE varies widely between individuals (Grelet et al., 2020; Shi et al., 2022), suggesting that NUE can potentially be improved by genetic selection with sufficient genetic variation. However, we found fewer studies on genetic selection for NUE in dairy cattle, except for the use of milk urea nitrogen (MUN) or milk urea (MU) implemented as proxies (Bastin et al., 2009; Bobbo et al., 2020). The reason for the lack of genetic studies of NUE, like the FE trait, is that it is a difficult and expensive phenotype to measure directly. So, the use of proxies of NUE to conduct genetic selection could be a good choice for its ease to be measured on a large scale.

2. Nitrogen metabolic processes and potential proxies of NUE in dairy cows

Before introducing the potential proxies of NUE, it is crucial to understand the metabolic processes of N in cows (Figure 1-1). Cows consume N mostly in the form of feed crude protein (CP), which includes rumen degradable protein and rumen undegradable protein (Aguirre-Villegas et al., 2017). The rumen degradable protein is degraded into different forms of N (e.g., amino acids and ammonia) by microorganisms in the rumen. The rumen undegradable protein will be partially digested in the small intestine, which becomes the direct source of amino acids for the cow. Ammonia is used for the *de novo* synthesis of amino acids by rumen microbes.

The excess of ammonia is, on the other hand, absorbed into the bloodstream through the rumen wall and converted to urea N in the liver, mostly discharged into milk, urine, and feces (Spek et al., 2013; Ariyarathne et al., 2019). In cattle, part of the urea synthesized by the liver also returns to the rumen through saliva and is reused by rumen microorganisms. Microbial proteins and undegraded proteins are broken down, in the small and large intestines by digestive enzymes coming from the digestive system, into amino acids. These amnio acids are absorbed and used by the cow's organism. The absorbed N is used for maintenance, reproduction, and milk production, whereas the unused N is eliminated from the body through urine and feces (Aguirre-Villegas et al., 2017). It is important to note that milk N was also divided into milk true protein N (MTPN) and MUN. Meanwhile, urea establishes a balance between various tissues and blood through cell membrane diffusion, resulting in a high correlation between blood urea N and MUN, which is the basis for predicting urinary N using MUN (Broderick and Clayton, 1997). We want to reduce N pollution from dairy cows by reducing MUN, although MUN can also be digested by humans.

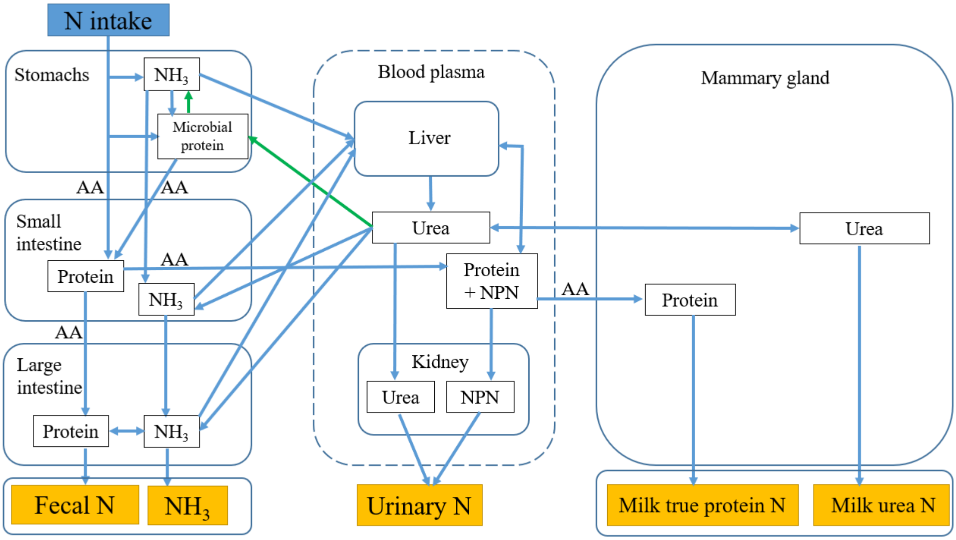


Figure 1-1. Metabolism of nitrogen in cattle (based on Spek et al., 2013). NPN is defined as non-protein-N except urea; AA is amino acids

Currently, there is only one N metabolism pathway (KEGG: 00910, https://www.genome.jp/pathway/map00910) in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The N metabolism pathway consists of six modules: N fixation (KEGG: M00175), Assimilatory nitrate reduction (KEGG: M00530), Dissimilatory nitrate reduction (KEGG: M00530), Denitrification (KEGG: M00529), Nitrification (KEGG: M00528), and Complete nitrification, comammox (KEGG: M00804). The identified metabolism pathway involves 18 genes *CA13*, *GLUD1*, *CA6*, *CA5A*, *CA9*, *GLUL*, *CA3*, *CPS1*, *CA8*, *CA2*, *CA4*, *CA12*, *CA5B*, *CA7*, *CA14*, *CA1*, *LOC784254*, and *LOC100847874* (https://biit.cs.ut.ee/gplink/l/s-vXmy-3RK). The genes containing CA are a synthetic series of carbonic anhydrases. Currently, no genome-wide association studies (GWAS) have been reported for NUE in dairy cattle, except for MU(N). However, none of the findings for MU(N) was to be significantly associated with any of the 18 genes mentioned above (Honerlagen et al., 2021; van den Berg et al., 2022a).

The general NUE of lactating dairy cows is defined as milk N divided by N intake (Calsamiglia et al., 2010). Milk N can be easily measured by using traits that are currently recorded (Milk N= Milk yield × Protein Percent / 6.38, WHO and FAO, 2011). For N intake calculations, we need to use individual cow feed intake and N concentration in her diet; however, measuring for both real traits is expensive and time-consuming. Therefore, the best solution is to establish an NUE proxy that can be easily implemented on the farm on a large scale and at a low cost. Lavery and Ferris (2021) summarized in detail the current methods and new strategies for the detection of NUE proxies and introduced NUE proxies with blood urea N, MUN, and N isotope analysis of cow tissues or fluids (plasma, milk, and hair). However, only predicted MUN or MU, as traditional NUE proxies have been reported in a large number of genetic analyses (Wood et al., 2003; Bastin et al., 2009; Bobbo et al., 2020), while other proxies (e.g., Predicted NUE (PNUE) and Predicted N losses (PNL)) have not been reported in a large number of studies. The main reason why MUN and MU are used as NUE proxies is that these traits have strong positive linear correlations with urinary N. However, the relationship between MUN and urinary N is influenced by various factors, such as dietary protein level, body weight, and water intake (Spek et al.,2013). Phenotypically, assuming urinary N is reduced, N intake potentially flows more to Milk N, whereas it is not clear whether genetic selection on MUN or MU indirectly can improve NUE.

The MUN also expressed as MU [MUN (mg/dL) = MU (mg/dL)/2.14] as reported by Beatson et al. (2019), is included as a standard part in most milk recording systems. The MU is usually predicted by milk MIR spectroscopy, which is rapid, cheap, and can be used in large-scale recording (Soyeurt et al., 2006; Gengler et al., 2016). The NUE proxies predicted by milk MIR are very promising applications in the genetic breeding area, and we have compiled the currently reported relevant studies (Table 1-1), except for MU. Ho et al. (2021) is a follow-up validation study for the study reported by Luke et al (2019) (Table 1-1). The current procedure of predicted blood (sera) urea N (BUN or SUN) is used to discern the metabolic status of the cow rather than to study NUE (Table 1-1). The PNUE throughout lactation can potentially be predicted by milk MIR, whereas PNL may only be predicted in early lactation (Grelet et al., 2020; Shi et al., 2022). However, it is important to note here that PNL is not really N pollution from cows. Meanwhile, multiple studies have shown that milk MIR can also be used to predict the DMI of cows (Lahart et al., 2019; Tedde et al., 2021). However, to calculate N intake, the protein percentage of the feed is required, which is difficult to popularize in grazing cattle. Therefore, relevant research is not given here. The basis for milk MIR's ability to predict N efficiency is as follows: the composition of milk is affected by the animal diet, and milk MIR spectra can reflect the changes in milk composition, which suggests that MIR can indirectly reflect the composition of the diet. (Klaffenböck et al., 2017). In addition to MUN (MU), we have introduced novel proxies of NUE.

It should also be mentioned that MUN yield (MUNY) is generally considered linearly proportional to the urinary urea-N excretion when defined as a quantity excreted (Wisconsin Alumni Research Foundation, 2015). So, we expect a decrease in N pollution by decreasing the MUNY. However, from the definition of NUE [(MTPN + MUNY) / N intake], lower MUNY may be associated with lower NUE. As presented in Figure 1-1, the MTPN and MUNY metabolic pathways are different, which suggests that we may simultaneously increase MTPN and decrease MUNY. By assuming a constant N intake, and increasing both MTPN and decreasing MUNY, we can obtain cattle with higher NUE and lower N pollution. Based on the above ideas, we proposed a new N efficiency index (NEI) (Figure 1-2) from a genetic point of view, as explained in Chapter 4 of this thesis.

Table 1-1. Prediction of NUE proxies using milk mid-infrared spectroscopy and its genetic analysis

Proxy[1]	Result[2]	Method[3]	N	Breed, Country[4]	DIM[5]	Reference	Genetic analysis
SUN	$R^2_V = 0.90$, $RMSE_V = 0.82$	PLS	878	HF, Australia	5-49	Luke et al., 2019	/
BUN	$R^2_V = 0.58$, $RPD_V = 1.55$	BiPLS	295	3 breeds, Italy	5-35	Benedet et al., 2019	Benedet et al., 2020[7]
SUN	$R^2_V = 0.69$, $RMSE_V = 0.17$	PLS	3,262	HF, Australia	0-50	Ho et al., 2021	van den Berg et al., 2021a, 2021b, 2022b
PNUE	$R^2_V = 0.68$, $RMSE_V = 5.01$	SVM	1,034	H, 3 country	4-50	Grelet et al., 2020	Chapter 3 of this thesis
PNUE	$R^2_V = 0.63$	SVM	600	H, China	154-452	Shi et al., 2022	/
PNL	$R^2_V = 0.59$, $RMSE_V = 0.07$	SVM	1,034	H, 3 country	4-50	Grelet et al., 2020	Chapter 3 of this thesis
PNL	$R^2_V = 0.19$	PLS	600	H, China	154-452	Shi et al., 2022	/

[1] SUN: Sera urea nitrogen; BUN: Blood urea nitrogen; PNUE: predicted nitrogen use efficiency; PNL: predicted nitrogen losses

[2] R^2_V: coefficient of determination in validation; $RMSE_V$: root mean square error of validation; RPD_V: ratio of performance to deviation in validation

[3] PLS: partial least squares regression; BiPLS: Backward interval partial least squares regression; SVM: support vector machines regression

[4] HF: Holstein Friesian; H: Holstein in the corresponding country, 3 breeds: Holstein Friesian, Brown Swiss, or Simmental; 3 countries: Denmark, Ireland, and the UK

[5] DIM: days in milk

[6] /: did not find related information

[7] Benedet et al., 2020: phenotype analysis using the equation from Benedet et al., 2019

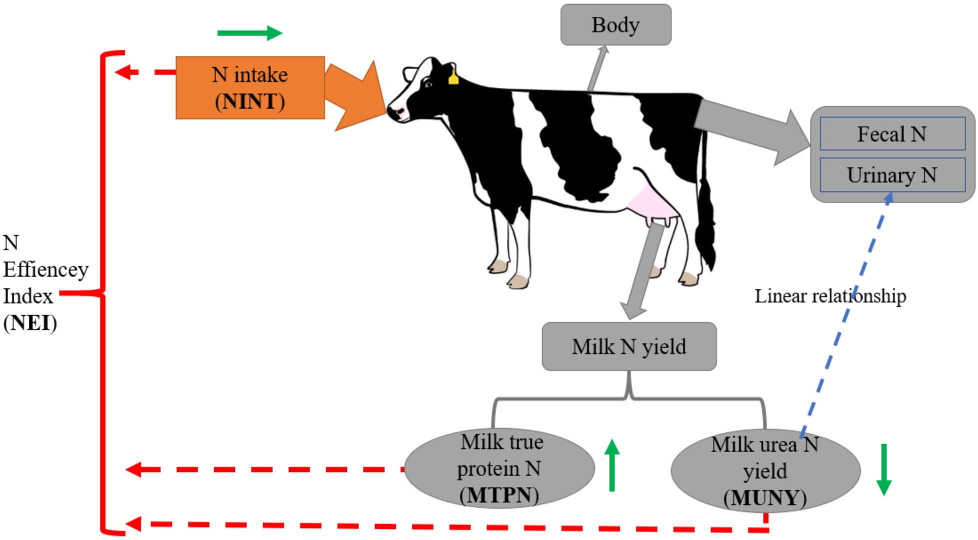


Figure 1-2. A new nitrogen efficiency index (NEI) for dairy cow

3. Research objectives and outline

The ultimate objective of this study was to explore the possibility of producing both more economical and more respectful of the environment cattle through genetic selection. Hence, the NUE of lactating dairy cows (Holstein in this thesis) was the target trait to be investigated. The hypothesis is that it is possible to simultaneously improve the NUE of cows and reduce N pollution through genetic selection. Such a high NUE herd would increase the economic interests of farmers while reducing N pollution and slowing climate changes, while meeting government policy requirements.

The objectives of this study were carried out in three parts (Figure 1-3). In the first part (Chapter 2), we focused on MU, the traditional NUE proxy, which is directly related to N pollution (urinary N) and is indirectly related to NUE. We performed a comprehensive genetic analysis of MU by using a multi-trait random regression model and estimated its approximate genetic correlations with 11 other traits of interest. In the second part (Chapter 3), the PNUE and PNL, novel proxies of NUE, were used for genetic analyses through a multi-trait repeatability model and their approximate genetic correlations with 30 other traits of interest were estimated. The PNUE and PNL of cows in early lactation were predicted by milk MIR (Grelet et al., 2020). In the third part (Chapters 4 to 6), a new NEI was established and the approximate genetic correlations between NEI and other 37 traits (indices) of interest were estimated (Chapter 4). Genomic data may help to have more accurate genomic predictions for low heritability traits, so genomic predictions for NEI and its composition traits were verified (Chapter 5). The genomic background of NEI and its composition traits was also analyzed through genome-wide association studies (GWAS) and post-GWAS (Chapter 6). Finally, a general discussion and a summary of the research findings were provided (Chapter 7).

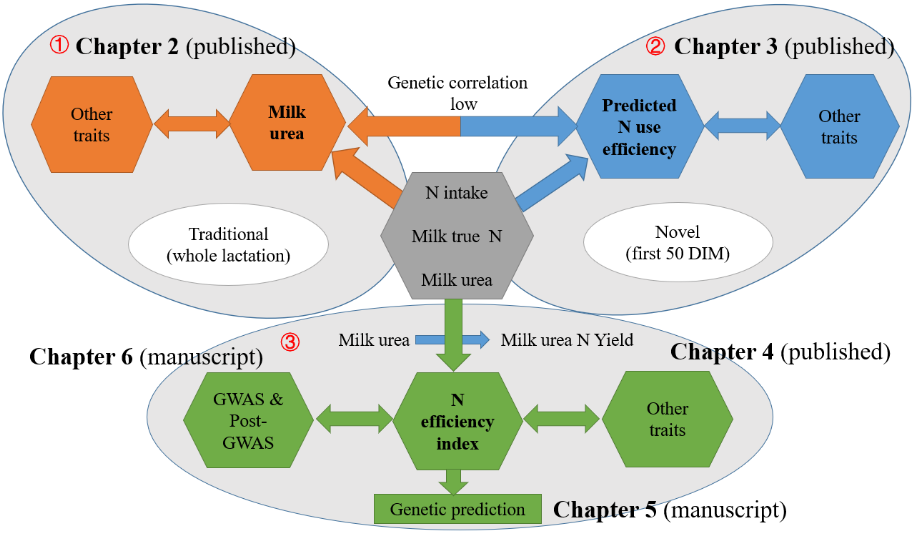


Figure 1-3. Outline of this thesis

4. References

aan den Toorn, S.I., E. Worrell, and M.A. van den Broek. 2021. How much can combinations of measures reduce methane and nitrous oxide emissions from European livestock husbandry and feed cultivation? J. Clean. Prod. 304:1–14. https://doi.org/10.1016/j.jclepro.2021.127138.

Adenuga, A.H., J. Davis, G. Hutchinson, T. Donnellan, and M. Patton. 2019. Environmental efficiency and pollution costs of nitrogen surplus in dairy farms: a parametric hyperbolic technology distance function approach. Environ. Resour. Econ. 74:1273–1298. https://doi.org/10.1007/s10640-019-00367-2.

Aguirre-Villegas, H.A., M.A. Wattiauxand, R.A. Larson, L. Chase, S.D. Ranathunga, and M.D. Ruark. 2017. Dairy cow nitrogen efficiency. University of Wisconsin-Extension. Accessed April 9, 2022. http://www.sustainabledairy.org/publications/Documents/Dairy%20Cow%20Nitrogen%20Efficiency%20A4131-09.pdf.

Bastin, C., L. Laloux, A. Gillon, F. Miglior, H. Soyeurt, H. Hammami, C. Bertozzi, and N. Gengler. 2009. Modeling milk urea of Walloon dairy cows in management perspectives. J. Dairy Sci. 92:3529–3540. https://doi.org/https://doi.org/10.3168/jds.2008-1904.

Benedet, A., M. Franzoi, C.L. Manuelian, M. Penasa, and M. De Marchi. 2020. Variation of blood metabolites of Brown swiss, Holstein-Friesian, and Simmental cows. Animals 10:1–12. https://doi.org/10.3390/ani10020271.

Benedet, A., M. Franzoi, M. Penasa, E. Pellattiero, and M. De Marchi. 2019. Prediction of blood metabolites from milk mid-infrared spectra in early-lactation cows. J. Dairy Sci. 102:11298–11307. https://doi.org/10.3168/jds.2019-16937.

Bobbo, T., M. Penasa, A. Rossoni, and M. Cassandro. 2020. Short communication: Genetic aspects of milk urea nitrogen and new indicators of nitrogen efficiency in dairy cows. J. Dairy Sci. 103:9207–9212. https://doi.org/10.3168/jds.2020-18445.

Brito, L.F., H.R. Oliveira, K. Houlahan, P.A.S. Fonseca, S. Lam, A.M. Butty, D.J. Seymour, G. Vargas, T.C.S. Chud, F.F. Silva, C.F. Baes, A. Cánovas, F. Miglior, and F.S. Schenkel. 2020. Genetic mechanisms underlying feed utilization and implementation of genomic selection for improved feed efficiency in dairy cattle. Can. J. Anim. Sci. 100:587–604. https://doi.org/10.1139/cjas-2019-0193.

Broderick, G.A., and M.K. Clayton. 1997. A statistical evaluation of animal and nutritional factors influencing concentrations of milk urea nitrogen. J. Dairy Sci. 80:2964–2971. https://doi.org/10.3168/jds.s0022-0302(97)76262-3.

Calsamiglia, S., A. Ferret, C.K. Reynolds, N.B. Kristensen, and A.M. Van Vuuren. 2010. Strategies for optimizing nitrogen use by ruminants. Animal 4:1184–1196. https://doi.org/10.1017/s1751731110000911.

Cavani, L., W.E. Brown, K.L. Parker Gaddis, R.J. Tempelman, M.J. VandeHaar, H. M. White, F. Peñagaricano, and K.A. Weigel. 2022. Estimates of genetic param eters for feeding behavior traits and its association with feed efficiency in Holst ein cows. Proceedings of the 12th World Congress on Genetics Applied to Live stock Production, Rotterdam, The Netherlands. Accessed July 12, 2022. https:// www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/14_016.pdf

de Ondarza, M.B., and J.M. Tricarico. 2017. REVIEW: Advantages and limitations of dairy efficiency measures and the effects of nutrition and feeding management interventions. Prof. Anim. Sci. 33:393–400. https://doi.org/10.15232/pas.2017-01624.

Gaddis, K., P. VanRaden, R. Tempelman, K. Weigel, H. White, F. Peñagaricano, J. Koltes, J. Santos, R. Baldwin, J. Burchard, J. Dürr, and M. VandeHaar. 2021. Implementation of feed saved evaluations in the U.S. Interbull Bull. 147–152.

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, and 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.

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 Consortim 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.

Ho, P.N., T.D.W. Luke, and J.E. Pryce. 2021. Validation of milk mid-infrared spectroscopy for predicting the metabolic status of lactating dairy cows in Australia. J. Dairy Sci. 104:4467–4477. https://doi.org/10.3168/jds.2020-19603.

Honerlagen, H., H. Reyer, M. Oster, S. Ponsuksili, N. Trakooljul, B. Kuhla, N. Reinsch, and K. Wimmers. 2021. Identification of genomic regions influencing N-Metabolism and N-Excretion in lactating Holstein-Friesians. Front. Genet. 12:1–12. https://doi.org/10.3389/fgene.2021.699550.

Lahart, B., S. McParland, E. Kennedy, T.M. Boland, T. Condon, M. Williams, N. Galvin, B. McCarthy, and F. Buckley. 2019. Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis. J. Dairy Sci. 102:8907–8918. https://doi.org/10.3168/jds.2019-16363.

Lavery, A., and C.P. Ferris. 2021. Proxy measures and novel strategies for estimating nitrogen utilisation efficiency in dairy cattle. Animals 11:1–15. https://doi.org/10.3390/ani11020343.

Luke, T.D.W., S. Rochfort, W.J. Wales, V. Bonfatti, L. Marett, and J.E. Pryce. 2019. Metabolic profiling of early-lactation dairy cows using milk mid-infrared spectra. J. Dairy Sci. 102:1747–1760. https://doi.org/10.3168/jds.2018-15103.

Madilindi, M.A., O.T. Zishiri, B. Dube, and C.B. Banga. 2022. Technological advances in genetic improvement of feed efficiency in dairy cattle: A review. Livest. Sci. 258:104871. https://doi.org/10.1016/j.livsci.2022.104871.

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.

Moran, D., and E. Wall. 2011. Livestock production and greenhouse gas emissions: Defining the problem and specifying solutions. Anim. Front. 1:19–25. https://doi.org/10.2527/af.2011-0012.

Reisinger, A., and H. Clark. 2018. How much do direct livestock emissions actually contribute to global warming? Glob. Chang. Biol. 24:1749–1761. https://doi.org/10.1111/gcb.13975.

Rochette, P., M.H. Chantigny, N. Ziadi, D.A. Angers, G. Bélanger, É. Charbonneau, D. Pellerin, C. Liang, and N. Bertrand. 2014. Soil nitrous oxide emissions after deposition of dairy cow excreta in eastern Canada. J. Environ. Qual. 43:829–841. https://doi.org/10.2134/jeq2013.11.0474.

Shi, R., W. Lou, B. Ducro, A.v.d. Linden, H.A. Mulder, S.J. Oosting, S. Li, and Y.

Wang. 2022. Predicting nitrogen use efficiency, nitrogen loss and dry matter intake of individual dairy cows in late lactation by including mid-infrared spectra of milk samples. Research Square. https://doi.org/10.21203/rs.3.rs-1722648/v1.

Siberski-Cooper, C.J., and J.E. Koltes. 2022. Opportunities to harness high-throughput and novel sensing phenotypes to improve feed efficiency in dairy cattle. Animals 12:1–17. https://doi.org/10.3390/ani12010015.

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.

Spanghero, M., and Z.M. Kowalski. 2021. Updating analysis of nitrogen balance experiments in dairy cows. J. Dairy Sci. 104:7725–7737. https://doi.org/10.3168/jds.2020-19656.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, and A. Bannink. 2013. A review of factors influencing milk urea concentration and its relationship with urinary urea excretion in lactating dairy cattle. J. Agric. Sci. 151:407–423. https://doi.org/10.1017/s0021859612000561.

Tedde, A., C. Grelet, P.N. Ho, J.E. Pryce, D. Hailemariam, Z. Wang, G. Plastow, N. Gengler, E. Froidmont, F. Dehareng, C. Bertozzi, M.A. Crowe, and H. Soyeurt. 2021. Multiple country approach to improve the test-day prediction of dairy cows' dry matter intake. Animals 11:1–19. https://doi.org/10.3390/ani11051316.

van den Berg, I., P.N. Ho, M. Haile-Mariam, P.R. Beatson, E. O'Connor, and J.E. Pryce. 2021a. Genetic parameters of blood urea nitrogen and milk urea nitrogen concentration in dairy cattle managed in pasture-based production systems of New Zealand and Australia. Anim. Prod. Sci. 61:1801–1810. https://doi.org/10.1071/an21049.

van den Berg, I., P.N. Ho, T.D.W. Luke, M. Haile-Mariam, S. Bolormaa, and J.E. Pryce. 2021b. The use of milk mid-infrared spectroscopy to improve genomic prediction accuracy of serum biomarkers. J. Dairy Sci. 104:2008–2017. https://doi.org/10.3168/jds.2020-19468.

van den Berg, I., P.N. Ho, T.V. Nguyen, M. Haile-Mariam, I.M. MacLeod, P.R. Beatson, E. O'Connor, and J.E. Pryce. 2022a. GWAS and genomic prediction of milk urea nitrogen in Australian and New Zealand dairy cattle. Genet. Sel. Evol. 54:1–12. https://doi.org/10.1186/s12711-022-00707-9.

van den Berg, I., P.N. Ho, T. V Nguyen, M.H. Mariam, T.D.W. Luke, and J.E. Pryce. 2022b. Using mid - infrared spectroscopy to increase GWAS power to detect QTL associated with blood urea nitrogen. Genet. Sel. Evol. 54:1–8. https://doi.org/10.1186/s12711-022-00719-5.

WHO and FAO (World Health Organization and Food and Agriculture Organization of the United Nations). 2011. Codex alimentarius: milk and milk products. 2nd ed. WHO FAO, Rome, Italy.

Wilcox, A.A., W.E. Carroll, R.E. Sterling, H.A. Davis, and A.G. Ware. 1966. Use of the Berthelot reaction in the the automated analysis of serum urea nitrogen. Clin. Chem. 12:151–157. https://doi.org/10.1093/clinchem/12.3.151.

Wisconsin Alumni Research Foundation. 2015. Milk Urea-N yield (MUNY) as a nutritional and environmental management tool for the dairy industry. US Pat. No. 9,955,673 B1.

Wood, G.M., P.J. Boettcher, J. Jamrozik, G.B. Jansen, and D.F. Kelton. 2003. Estimation of genetic parameters for concentrations of milk urea nitrogen. J. Dairy Sci. 86:2462–2469. https://doi.org/10.3168/jds.s0022-0302(03)73840-5.

Wood, G.M., P.J. Boettcher, J. Jamrozik, G.B. Jansen, and D.F. Kelton. 2003. Estimation of genetic parameters for concentrations of milk urea nitrogen. J. Dairy Sci. 86:2462–2469. https://doi.org/10.3168/jds.s0022-0302(03)73840-5.

2

Chapter II Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows

Adapted from: Chen, Y., H. Atashi, S. Vanderick, R.R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2021. Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows. J. Dairy Sci. 104:12741-12755. https://doi.org/10.3168/jds.2021-20659.

Foreword

Milk urea concentration (MU) is a traditional proxy to monitor the nutritional status of dairy cows and reduce nitrogen emissions into the environment. The reason is that MU and urinary nitrogen are often considered to have a linear relationship. This thesis begins with a genetic analysis of MU, exploring the possibility and potential impact of adding a proxy traditional of nitrogen use efficiency to the genetic selection system in the Walloon Region of Belgium.

The models used in this Chapter were nearly identical to the current genetic evaluation model used for yield traits in the Walloon Region of Belgium. However, these models can be further optimized, such as by exploring the use of different orders of polynomials.

Abstract

This study aimed to estimate the genetic parameters of milk urea concentration (MU) and its genetic correlations with milk production traits, longevity, and functional traits in the first 3 parities in dairy cows. The edited data set consisted of 9,107,349 MU test-day records from the first 3 parities of 560,739 cows in 2,356 herds collected during the years 1994 to 2020. To estimate the genetic parameters of MU, data of 109 randomly selected herds, with a total of 770,016 MU test-day records, were used. Genetic parameters and estimated breeding values were estimated using a multiple-trait (parity) random regression model. Herd-test-day, age-year-season of calving, and days in milk classes (every 5 d as a class) were used as fixed effects, whereas effects of herd-year of calving, permanent environment, and animal were modeled using random regressions and Legendre polynomials of order 2. The average daily heritability and repeatability of MU during days in milk 5 to 365 in the first 3 parities were 0.19, 0.22, 0.20, and 0.48, 0.48, 0.47, respectively. The mean genetic correlation estimated among MU in the first 3 parities ranged from 0.96 to 0.97. The average daily estimated breeding values for MU of the selected bulls (n = 1,900) ranged from −9.09 to 7.37 mg/dL. In the last 10 yr, the genetic trend of MU has gradually increased. The genetic correlation between MU and 11 traits of interest ranged from –0.28 (milk yield) to 0.28 (somatic cell score). The findings of this study can be used as the first step for the development of a routine genetic evaluation for MU and its inclusion into the genetic selection program in the Walloon Region of Belgium.

Key words: heritability, genetic correlation, random regression, milk urea

1. Introduction

Milk urea nitrogen, also expressed as milk urea concentration (MU) [MUN (mg/dL) = MU (mg/dL)/2.14] as reported by Beatson et al. (2019), is included as a standard part in most milk recording systems. Regardless of the expression used, MUN or MU are indicators used to monitor the nutritional status of dairy cows and reduce nitrogen emissions to the environment (Samoré et al., 2007; Spek et al., 2013). Feed proteins are decomposed by rumen microorganisms into ammonia and branched-chain fatty acids. The ammonia is then used for de novo synthesis of amino acids by rumen microbes, whereas excess ammonia is absorbed into the bloodstream through the rumen wall and converted to urea nitrogen in the liver, which is then discharged into milk, urine, and feces (Spek et al., 2013; Ariyarathne et al., 2019). Among them, urine and feces urea N contribute to water pollution and gaseous N emissions (Hristov et al., 2019). Although it is difficult to assess urine urea N in routine dairy farming operations, MUN is a normal milk component (Gengler et al., 2016) that can be easily measured (Beatson et al., 2019). Even if there is not a total consensus, previous studies (e.g., Jonker et al., 1998) considered MUN as a valid predictor of urine urea N excretion. Kauffman and St-Pierre (2001) reported a linear relationship between urinary N and MUN in dairy cows. Some studies showed that MUN can reflect the efficiency of protein metabolism of dairy cows (e.g., Bastin et al., 2009), and the balance between CP and energy in the diet (e.g., Roy et al., 2003).

The MU varies between individuals and is affected by many factors. Literature reported among others, the factors of feed compositions (Correa-Luna et al., 2019), milking time (Bendelja et al., 2011), age at calving (Wood et al., 2003), stage of lactation and season of milking (Fatehi et al., 2012; Yin et al., 2012), herd and parity (Čobanović et al., 2017; Siatka et al., 2020), and genetics (Bendelja et al., 2011; Mucha and Strandberg, 2011; Rzewuska and Strabel, 2013).

Inclusion of a given trait into a breeding program is based on the identification of its potential contribution to the breeding goal. For MU, its potential contribution can be explained as follows. First, some researchers demonstrated the direct interest in breeding for lower MU, or MUN, to reduce the environmental effect (Bobbo et al., 2020; Marshall et al., 2020). Lower MU (or MUN) could also be linked to improved N use efficiency but under the hypothesis that reduced MU (or MUN) does not affect the amount of protein produced (Jonker et al., 1998). In addition, to reduce potential unfavorable correlated responses, correlations of MU and traits included in the breeding goal have to be considered. Previous research showed that MU (or MUN) is correlated with yield traits (e.g., milk, protein, and fat yields; Yin et al., 2012; Beatson et al., 2019; Ariyarathne et al., 2021), type traits (Satoła et al., 2017b), BW (Hojman et al., 2005), SCC (Bobbo et al., 2019), and reproductive performance (Siatka et al., 2020; Kananub et al., 2020). However, the results did not all indicate similar direction and strength of correlations. Therefore, more studies are needed to make the right decisions about the position of MU in the future breeding programs in the Walloon Region of Belgium.

Although MU has been included in the regular milk recording of dairy cows in the Walloon Region of Belgium since 1994, its use is restricted to management recommendations. In this context, Bastin et al. (2009) used advanced modeling (e.g., Mayeres et al., 2004) of test-day MU records in the first parity in the Walloon Region. Their study included also genetic effects but did not use EBV and did not investigate genetic correlations of MU with other economically important traits. Therefore, the present study can be considered also being a complement and follow-up of this previous research in the same environment but more than 10 yr ago.

This study had 3 purposes: (1) estimating the genetic parameters of MU in the first 3 parities of Walloon dairy cows using a random regression animal model, (2) using the estimated parameters to perform a genetic evaluation for MU, and (3) estimating approximate genetic correlations between MU and 11 traits of interest currently evaluated in the Walloon Region of Belgium.

2. Materials and methods

2.1. Data

Following previous research reported by Bastin et al. (2009) and the usual trait definition in the Walloon Region of Belgium, expression as MU was preferred to MUN. The MU records were generated by the official milk recording in the Walloon Region of Belgium using mid-infrared spectrometry and commercially available instruments and calibrations from FOSS (Foss Electric A/S). Data used in this study were extracted from the generated database to create a genetic evaluation data set. Records from only the first 3 lactations were kept. Hereafter we will call the MU in the first 3 lactations MU1, MU2, and MU3, respectively. Further standard edits made were the same as for routine genetic evaluations for yield traits. Therefore, records from DIM lower than 5 and greater than 365 d were excluded. The value for MU was restricted between 2 and 70 mg/dL. Finally, the 9,107,349 test-day records of the first 3 parities on 560,739 cows in 2,356 herds collected from 1994 to 2020 were kept. More than 95% of cows included were Holstein. Pedigree data were extracted from the database used for the official Walloon genetic evaluations and contained 814,023 animals (43,162 males and 770,861 females).

2.2. (Co)variance Component Estimation

Data Set. The size of the original data set and the expected complexity of the used model required that (co)variance component estimation (VCE) was done on subsamples extracted from the whole genetic evaluation data set. We decided to repeat VCE 10 times, requiring 10 independent subsamples. Sampling was based on herds. Even if we identified a large number of herds in the data (2,356), due to the specificity of Walloon data, many herds did not record milk over a longer period, or stopped and started milk recording over the years. Therefore, a threshold of at least 5,000 available test-day records taken over the considered period of time was defined. These earmarked 629 herds are to be available for the next step. Ten times, herds were randomly drawn from the pool of 629 herds until 3,000 cows with records were obtained. This was achieved for 9 subsamples with 11 herds, and in one case with 10 herds, leading to 10 subsets of a total of 109 herds randomly drawn from the 629 herds. To create independent subsamples, herds could only be drawn once. Moreover, calving age for the first, second, and third parity was restricted to 22 to 37, 34 to 51, and 46 to 65 mo, respectively.

Model. Basic models used for VCE and genetic evaluation were nearly identical and based on the current genetic evaluation model used for yield traits in the Walloon Region of Belgium (Croquet et al., 2006). The following multiple-trait random regression model was used where records in each lactation were considered as separate traits:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Xb} + \mathbf{Dd} + \mathbf{Q}(\mathbf{Cc} + \mathbf{Wp} + \mathbf{Za}) + \mathbf{e}, \qquad [1]$$

where **y** is the vector of MU records in the first 3 lactations (traits MU1, MU2, and MU3), **h** is the vector of fixed herd × test-day classes, and **b** is the vector of fixed age-year-season of calving classes (4 seasons were de-fined: December to February, March to May, June to August, and September to November), **d** is the vector of fixed lactation stage classes (72 classes, DIM was divided by into 5-d classes, except from DIM 360 to 365 which was considered as one class), **c** is the vector of herd × year of calving common environmental random regression coefficients, **p** is the vector of permanent environmental random regression coefficients, **a** is the vector of additive genetic random regression coefficients, **e** is the vector of random residuals effects, **H**, **X**, **D**, **C**, **W**, and **Z** are incidence matrices assigning observations to the corresponding effects, **Q** is the covariate matrix for second-order modified Legendre polynomials associated with DIM d with

$$q_{0(d)} = 1.0,$$

$$q_{1(d)} = 3.0^{0.5}x,$$

$$q_{2(d)} = \left(\frac{5}{4}\right)^{0.5}(3.0x^2 - 1),$$

where $x = -1 + 2\left(\frac{d-1}{365-1}\right)$.

The modified Legendre polynomials referred to the formula proposed by Gengler et al. (1999), but the DIM range was changed following the current practice in the routine genetic evaluation in the Walloon Region of Belgium.

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

$$\mathrm{E}(\mathbf{y}) = \mathbf{Hh} + \mathbf{Xb} + \mathbf{Dd},$$

$$\mathrm{E}(\mathbf{c}) = \mathrm{E}(\mathbf{p}) = \mathrm{E}(\mathbf{a}) = \mathrm{E}(\mathbf{e}) = 0,$$

The following (co)variance structures were assumed:

$$\mathrm{Var}\begin{bmatrix}\mathbf{c}\\ \mathbf{p}\\ \mathbf{a}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{C}^{*}\otimes\mathbf{I}_{c} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{P}\otimes\mathbf{I}_{w} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{G}\otimes\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}\end{bmatrix}$$

where $\mathbf{C}^{*}$ contained the coefficients of the herd × year of calving (co)variance function (9 × 9 matrix); **P** contained the coefficients of the permanent environment (co)variance function (9 × 9 matrix); **G** contained the coefficients of the genetic (co)variance function (9 × 9 matrix); $\mathbf{I_c}$was an identity matrix of dimension c (number of herd × year of calving classes); $\mathbf{I}_w$ was an identity matrix of dimension w (number of cows with records); **A** was the additive genetic relationship matrix among all animals in the pedigree; **R** was a diagonal matrix of dimension n (total number of MU records across the 3 lactations) with diagonal elements equal to $\sigma^2_{e(t)}$ which was the residual variance for trait t (lactation) in which MU was recorded (MU1, MU2, MU3); and ⊗ represents the Kronecker product function.

Computations were done using the BLUPF90 family of programs (Misztal et al., 2018). Variance components were estimated using REML by EM algorithm as implemented in REMLF90 (version 1.84) program.

After estimating the (co)variance components separately for each subset, values were averaged and the standard deviation (SD) of each (co)variance component was computed as the approximate sampling error. Homogeneity of residual variance was checked visually by computing and plotting the SD of observed residuals (difference between observed and predicted values) for each class of DIM in the first 3 parities.

2.3. Genetic Parameter Calculation

Based on estimated (co)variance components, several genetic parameters were computed. For each trait of MU1, MU2, MU3 (t), heritability (h^2_{td}) and repeatability (rep_{td}) were estimated for individual DIM (d) between 5 and 365 as follows:

$$\frac{\sigma^2_{a(td)}}{\left(\sigma^2_{a(td)} + \sigma^2_{p(td)} + \sigma^2_{c(td)} + \sigma^2_{e(t)}\right)} \text{ and}$$

$$\frac{(\sigma^2_{a(td)} + \sigma^2_{p(td)} + \sigma^2_{c(td)})}{(\sigma^2_{a(td)} + \sigma^2_{p(td)} + \sigma^2_{c(td)} + \sigma^2_{e(t)})},$$

Genetic (co)variance for average 305-d MU was obtained by using $\mathbf{G}_{\mathbf{lact}} = \mathbf{SGS'}$, where $\mathbf{G}_{\mathbf{lact}}$ was the 3 × 3 (co)variance matrices among average 305-d lactation MU for the three traits (MU1, MU2, and MU3); **G** was a covariance matrix (9 × 9) of genetic additive effect; and **S** was a 3 × 9 summation matrix that averaged the contributions of a given test-day to the 305-d MU for the three traits (MU1, MU2, and MU3). The same approach was used to derive $\mathbf{C}_{\mathbf{lact}}$ and $\mathbf{P}_{\mathbf{lact}}$ matrices, which represent the herd × year of calving and permanent environmental (co)variances for average 305-d MU. Heritability for average 305-d MU was computed as the ratio of the genetic variance to the sum of the genetic, permanent environmental, herd × year of calving, and residual variances. Correlations for average 305-d MU between parity i and j were computed as the ratio of the average covariance-305 cov(i, j) to the square root of the products of the average variances-305 of MU in parity i and j. The residual variance was the same for the whole lactation in each parity.

2.4. Genetic Evaluation

The model used for calculating EBV was the same as that used for (co)variances estimation. Breeding values for the included animals were estimated according to the algorithm implemented in the BLUP90IOD2 (version 3.81; http://nce.ads.uga.edu/wiki/doku.php?id=readme.pcg2) program. Preparation of different types of EBV was done following the example of yield traits in the routine evaluations. Daily EBV was calculated using the following equation: $\mathrm{EBV}_{ktd} = \mathbf{q}'_{m(d)}\mathbf{a}_{ktm}$, where EBV_{ktd} was the EBV of cow k for trait t (here MU1, MU2, and MU3) in DIM of d, for DIM 5 to 365; $\mathbf{a}_{ktm}$ were the 3 solutions m (additive genetic random regression coefficients) of cow k for trait t by BLUP90IOD2; $\mathbf{q}_{m(d)}$ was the same as the **Q** in the equation (1), in which d ranged from 5 to 365. Following the example of reported EBV for yield traits that are average lactation EBV of the first three parities, the average daily EBV value of MU in the first three parities was also calculated. The formula for average daily EBV was:

$$\mathrm{EBV}_k = \sum_{j=1}^{3} \sum_{d=1}^{305} \mathbf{q}'_{mj(d)} \mathbf{a}_{ktmj} / (3 \times 305),$$

where *j* is the number of parity, and the rest parameters are the same as the formula of daily EBV. The reliability (REL) of average daily EBV for MU was calculated based on the INTERBULL calculation method proposed by Stranden et al. (2000) as used in the Walloon routine genetic evaluation system for yield traits. This system provides also directly expected daughter contributions (EDC) as defined by INTERBULL. In the subsequent analysis of approximate genetic correlation with traits included, the MU in the first three parities were combined into one trait. The same idea was also recommended by Hossein-Zadeh and Ardalan (2011).

2.5. Approximate Genetic Correlations

The approximate genetic correlations between MU and 11 traits of interest were estimated using the procedure proposed by Blanchard et al. (1983) which is a generalized version of Calo's method (Calo et al., 1973) when selected bulls had very diverse REL. The 11 traits of interest included milk yield (MY), fat yield (FY), protein yield (PY), SCS, fat percentage (FP), protein percentage (PP), longevity, female fertility (FF), BCS, direct calving ease (DCE), and maternal calving ease (MCE). The calculation model and algorithm of EBV for the 11 traits of interest can be found in the documents (Belgium) submitted to Interbull (https://interbull.org/ib/geforms) or reports by Vanderick et al. (2020). To use those bulls that were commonly used in Walloon, in the first step bulls had to have at least 100 descendants registered. Then, a total of 1,900 bulls were identified as having REL for average daily EBV of MU that was greater than 0.50 associated with at least 5 daughters with MU records. Based on the data obtained for the traits included, 2 different strategies were used to calculate the approximate genetic correlations.

2.6. Correlations Based on Daily EBV of MU, Yield Traits, and SCS

Even if they are not provided in routine, daily EBV of MY, FY, PY, and SCS (4 traits) for each DIM can be generated using the genetic evaluation system of dairy cows in the Walloon Region of Belgium. In this study, we used those EBV obtained during the routine run of August 2020. As REL of daily EBV are not calculated in routine, the following procedure was used to get reasonable approximated REL associated with daily EBV. First lactation based REL (REL_{305}) were transformed to

EDC using the lactation heritability for each trait (h^2_{305}), *EDC* representing the information content:

$$EDC = \frac{REL_{305}}{1 - REL_{305}} \times \frac{4 - h^2_{305}}{h^2_{305}}$$

In the second step, the *EDC* was back transformed to Daily-REL:

$$REL_d = \frac{EDC}{EDC + \frac{4 - h^2_d}{h^2_d}}$$

where REL_d and h^2_d were, respectively, the REL and heritability at DIM d for each trait. The needed h^2_d was calculated using the official (co)variance matrix of the corresponding traits. Based on the EBV of the studied traits and the corresponding REL in the given test day, the approximate genetic correlations between MU and MY, FY, PY, and SCS were calculated, for DIM 5 to 365.

2.7. Correlations based on Selected daily EBV of MU and EBV of Fat and Protein Percentages Longevity, and Selected Functional Traits

For 7 traits (FP, PP, Longevity, FF, BCS, MCE, and DCE) evaluated in Walloon that are not defined across a longitudinal lactation scale, the following method was used to calculate their approximated genetic correlations with MU at different lactation stages (25, 125, 225, and 325 DIM). The EBV and REL of the 7 traits examined for 1,900 previously selected bulls were published on the ELINFO website (https://www.elinfo.be/telechargerEN.html). First among the 1,900 previously selected bulls, based on reliabilities for the 7 traits examined with REL greater than 0.50, a total of 1,317 bulls were selected. Then, based on the Daily-EBV of MU at selected DIM (25, 125, 225, and 325). and their REL, the EBV of the 7 traits examined, and their corresponding REL, the approximate genetic correlations between them were calculated. Standard errors of the estimated approximate genetic correlation were calculated using 1,000 bootstrapped replicates (Chen et al., 2021).

In addition, the genetic trends were analyzed by year of birth for all 12 traits included based on the selected bulls. For genetic trend analysis of MU, the average Daily-EBV of selected bulls (n = 1,900) was used; while for the other 11 traits examined the EBV of selected bulls (n = 1,900 or n = 1,317) from the ELINFO website were used. The EBV of all traits had been standardized (mean = 0, SD = 1) because the values vary greatly. According to the birth year, the bulls were divided into the following 7 groups: < 1992, 1992-1995, 1996-1999, 2000-2003, 2004-2007, 2008-2011, and 2012-2015. The number of bulls in each group was more than 100. The data preparation and processing were done using R (R Core Team, 2020).

3. Results and discussion

3.1. Descriptive Statistics

The description of the final edited data set (n = 9,107,349) used for genetic evaluation and the sampled data sets for VCE (n = 770,016) is presented in Table 2-1. The average and SD of MU for the sampled VCE data sets were slightly lower than those found for the final edited data set. The average MU in each of the first 3 lactations ranged from 24.80 to 26.19 mg/dL, which were similar to those reported by Bastin et al. (2009) for the first-parity average MU (25.51 mg/dL) of dairy cows in Walloon. These values were also consistent with those reported in previous studies (Čobanović et al., 2017; Lopez-Villalobos et al., 2018; Ariyarathne et al., 2019). However, Satoła et al. (2017b) reported a lower value for MU in Polish Holstein-Friesian cows. The differences observed for average MU can partly be attributed to different feeding management and the studied populations. The average value of MU found in this study is within the range (15.0 to 30.0 mg/dL) recommended for cattle management in Germany, France, and Austria (Glatz-Hoppe et al., 2020). The coefficient of variation (CV%) of MU ranged from 35.08 to 40.80%with a mean of 40.51%, which is slightly lower than that (42%) reported by Bastin et al. (2009).

Table 2-1. Descriptive statistics of milk urea concentration (mg/dL) of the final edited and the used data set

Parity	No. of records	No.of cows	Minimum	Maximum	Mean	SD	CV(%)
Genetic evaluation dataset							
1	3,933,442	500,352	2	70	26.01	10.40	39.99
2	3,020,958	387,040	2	70	26.19	10.67	40.73
3	2,142,949	276,624	2	70	26.08	10.64	40.80
All	9,107,349	560,739	2	70	26.10	10.57	40.51
(Co)variance component estimation dataset							
1	261,906	29,548	2	70	25.70	9.29	36.15
2	265,709	29,389	2	70	25.30	9.11	36.01
3	242,401	29,449	2	70	24.80	8.70	35.08
All	770,016	30,090	2	70	25.28	9.03	35.74

Figure 2-1 shows the curves of average MU over DIM classes for the first 3 parities. The value of MU was slightly higher in the first few weeks (lowest value around DIM 23), but it quickly increased to higher levels. However, at the end of the lactation, a decreasing trend was found especially for multiparous cows. The lactation curves observed for MU during the first 305 DIM were the same as those reported by Wood et al. (2003) and Bastin et al. (2009); however, in the extended part of the lactation (i.e., after 305 DIM), the trend showed a gradual decrease. In comparison, based on a limited number of records (n = 5,576), Stoop et al. (2007) reported an increase followed by a decreasing trend for MU. Their curve was similar to the lactation curve for milk yield but with a high level of MU maintained post-lactation peak. The MU reached its lowest value in the early lactation, indicating that dairy cows have a higher N use efficiency at this time. It is well documented that dairy cows are in a negative energy balance state in the early stage of lactation because the feed intake of dairy cows cannot meet their requirements. During this time, it can be assumed that cows use N as efficiently as possible which can explain, at least in part, the obtained results.

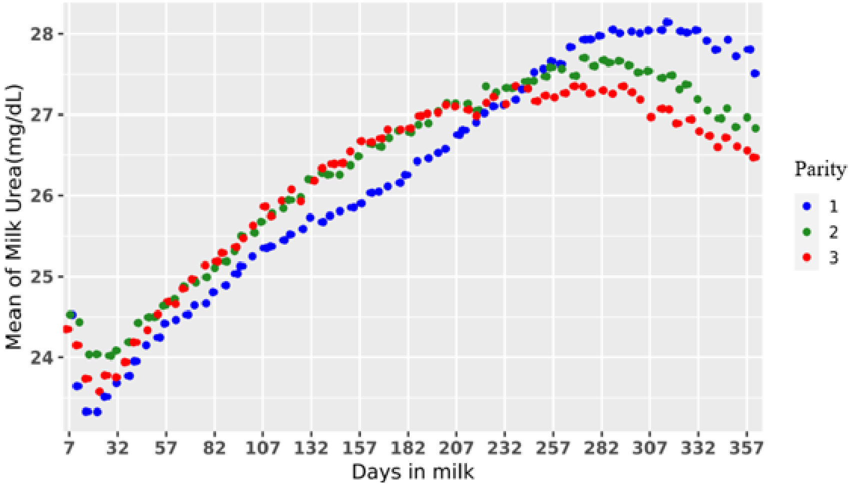


Figure 2-1. Lactation curves for milk urea concentration within classes of DIM (represented by their average DIM) for the first 3 parities (n = 9,107,349)

3.2. Genetic Parameters

The estimated variances for all random effects of MU were for all parities higher at the beginning and the end of the lactation, lower in the middle, and reached their maximum levels at 365 DIM (Figure 2-2). Figure 2-2C shows that the additive genetic variance of MU reached its minimum level at around 50 DIM, whereas its maximum value was found at DIM 365. Our results were in agreement with those reported previously (Bastin et al., 2009; Rzewuska and Strabel, 2013; Ariyarathne et al., 2019). The addition of the herd × year of calving random effect in the model helped to avoid an excessive border effect of the additive genetic variance exceeding 305 DIM (Figure 2-2A).

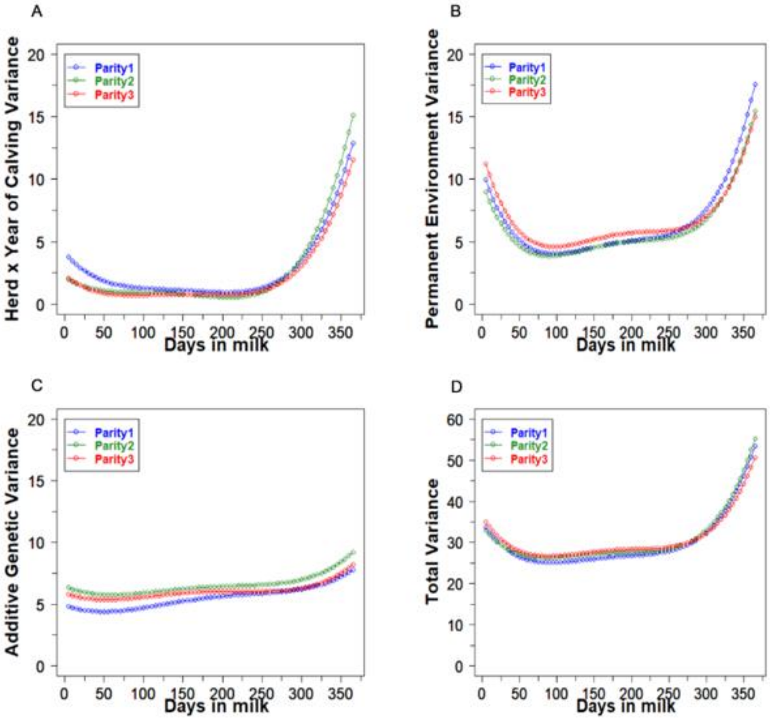


Figure 2-2. Variance components estimated for milk urea concentration over DIM for the first 3 parities

Table 2-2. Mean (approximate SE computed from SD of the 10 sampling subsets) of daily heritability (h^2), repeatability, herd × year of calving variance (σ^2_c), permanent environmental variance (σ^2_p), additive genetic variance (σ^2_a), and total variance (σ^2_{total}) of milk urea concentration for each parity throughout the lactation

Parity	h^2	Repeatability	σ^2_c	σ^2_p	σ^2_a	σ^2_{total}
1	0.19 ± 0.02	0.48 ± 0.08	2.70 ± 2.73	6.57 ± 3.06	5.52 ± 0.86	29.92 ± 6.41
2	0.22 ± 0.02	0.48 ± 0.08	2.51 ± 3.36	6.07 ± 2.56	6.55 ± 0.78	30.69 ± 6.58
3	0.20 ± 0.02	0.47 ± 0.07	2.11 ± 2.52	6.72 ± 2.35	6.05 ± 0.61	30.73 ± 5.30

The estimated variance components for MU in the first 3 parities averaged over the lactation period (DIM 5–365) along with their SD are presented in Table 2-2. The additive genetic variances for the first, second, and third parity accounted for 18.45, 21.34, and 19.69% of the total phenotypic variance, respectively. Furthermore, the calculated SD of the observed residuals accumulated across the 10 VCE subsets for each DIM class were nearly horizontal except for the beginning (first 15 DIM) and the end (last 30 DIM) of the lactation (Figure 2-3). The higher residual variance observed at the beginning of lactation could also be attributed to larger changes in body metabolism during this period (e.g., negative energy balance) that could not be perfectly described by our model. There are at least 2 potential explanations for the decrease at the end of the lactation. First, the observed environmental variances tended to be very high at the end of the lactations, the slight decrease could be compensation. The lower residual variance observed at the end of lactation may also be due to the decrease in the amount of data available in this period (half compared with the previous classes). As illustrated in Figure 2-3, based on the visual inspection, despite these very slight deviations, the overall horizontal nature of the residual SD curves can be confirmed. This strengthens the assumption that modeling homogeneous residuals across the lactation is a valid assumption that simplified the used model. Especially for genetic evaluations, a simpler model is advantageous in avoiding preadjustment or weighting records to correct for heteroscedasticity. The average (SD) residual variances across the 10 VCE subsets for MU1, MU2, and MU3 were 15.13 (0.87), 15.56 (0.92), and 15.86 (1.01), respectively.

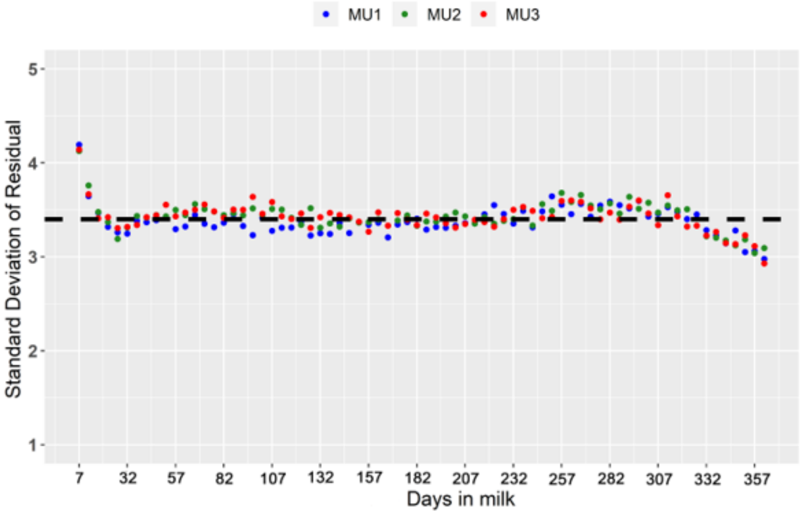


Figure 2-3. The standard deviation (SD) of the observed residual within classes of DIM (represented by their average DIM) for milk urea concentration of the first 3 parities in the sampled data set (n = 770,016) used for (co)variance component estimation (MUx: milk urea concentration in parity x, where x is 1, 2, or 3)

The heritability and repeatability estimate of MU by DIM are presented in Figure 2-4. The heritability estimates showed an increasing phase from the beginning of the lactation, reached its maximum level at around 215 DIM, then decreased to the end of the lactation (from 0.13 to 0.23), which is in a close agreement with that reported by Yin et al. (2012). However, the trend of heritability reported by Rzewuska and Strabel (2013) was different, which may be, at least in part, due to different models used or different populations studied. The MU heritabilities averaged across lactation were 0.19, 0.22, and 0.20, for the first, second, and third parity, respectively (Table 2-2), which are in a close agreement with those reported by Mitchell et al. (2005), Mucha and Strandberg (2011), and Lopez-Villalobos et al. (2018). However, these values were higher than those reported by König et al. (2008) and Bastin et al. (2009), and lower than those reported by Miglior et al. (2007) and Ariyarathne et al. (2019). The average 305-d heritabilities of MU were 0.52, 0.60, and 0.54, respectively, in the first 3 parities (Table 2-3), which were similar to the maximum average heritability of MUN (0.59 in the second parity) reported by Wood et al. (2003). The trend of repeatability for MU over DIM resembled a “U” shape (from 0.40 to 0.72) and reached its maximum at 365 DIM, which is in agreement with Ariyarathne et al. (2021). The repeatability estimates of MU averaged over lactation were 0.48, 0.48, and 0.47, respectively, for the first, second, and third parities (Table 2-2), which are in line with those reported in previous studies (Ariyarathne et al., 2019, 2021; Beatson et al., 2019).

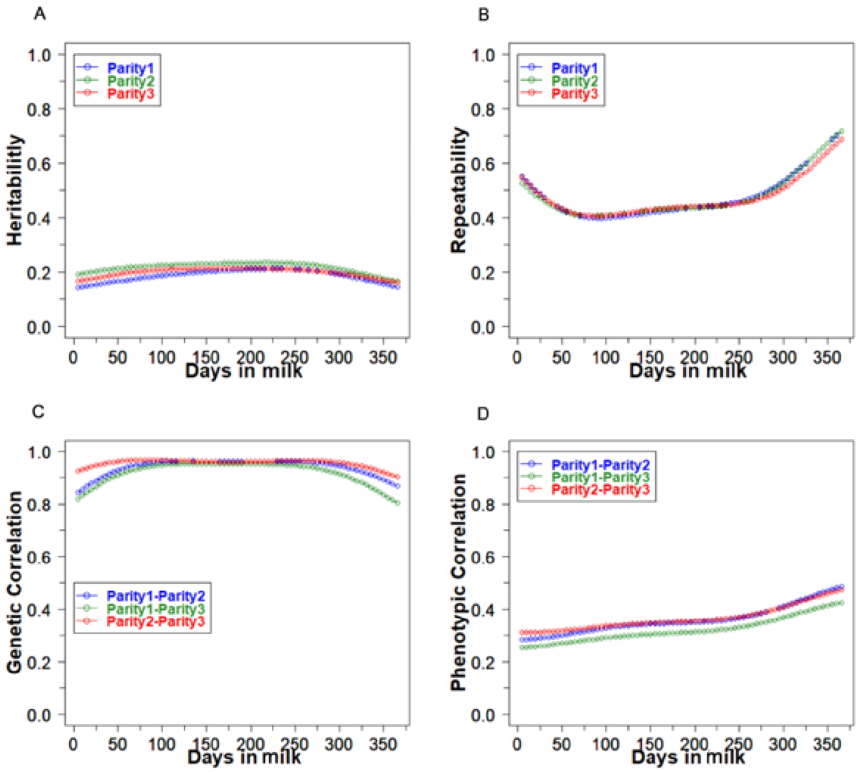


Figure 2-4. Heritability (A), repeatability (B), genetic correlation (C), and phenotypic correlation (D) of milk urea concentration over DIM in the first 3 parities

3.3. Correlations Among the First 3 Parities

The genetic and phenotypic correlations of MU among the first 3 parities are presented in Figure 2-4C and D, respectively. The genetic correlations of MU among the first 3 parities were lower at the beginning and at the end of the lactation, whereas their maximum levels were found in the middle of the lactation. The average genetic correlations of MU calculated among the first 3 parities during DIM 5 to 365 ranged from 0.92 to 0.96. The obtained results are similar to those reported by Hossein-Zadeh and Ardalan (2011) and Rzewuska and Strabel (2013), but higher than those reported by Satoła et al. (2017a). The 305-d genetic correlation of MU among the first 3 parities ranged from 0.96 to 0.97 (Table 2-3), which is in a close agreement with that found between the first and second parity by Mitchell et al. (2005).

Table 2-3. Average (± SE) 305-d heritabilities (diagonal), average 305-d genetic correlations (above diagonal), and average 305-d phenotypic correlations (below diagonal) of milk urea concentration among the first 3 parities in Walloon dairy cows computed from the 10 sampling subsets (approximate SE computed from observed SD)

Parity	1	2	3
1	0.52 ± 0.08	0.97 ± 0.01	0.96 ± 0.02
2	0.84 ± 0.03	0.60 ± 0.07	0.97 ± 0.07
3	0.76 ± 0.03	0.85 ± 0.02	0.54 ± 0.07

The phenotypic correlations of MU among the first 3 parities increased with increasing DIM and ranged from 0.32 to 0.37 in agreement with those reported by Hossein-Zadeh and Ardalan (2011). The 305-d phenotypic correlations of MU among the first 3 parities were relatively high and ranged from 0.76 to 0.85 (Table 2-3).

3.4. Genetic Evaluation and Trend

The Figure 2-5 showed the distribution of standardized average daily EBV of MU and their associated REL for the selected bulls (mean = 0, SD = 1). The average daily EBV of 1,900 bulls and the corresponding REL ranged from −9.09 to 7.37 mg/dL and 0.50 to 0.99 (mean = 0.90), respectively. The relatively high REL found for average daily EBV may be attributed to the high 305-d heritability of MU. In addition, only bulls with at least 100 daughters in the pedigree were selected in this study which also can explain the relatively high REL found for average daily EBV.

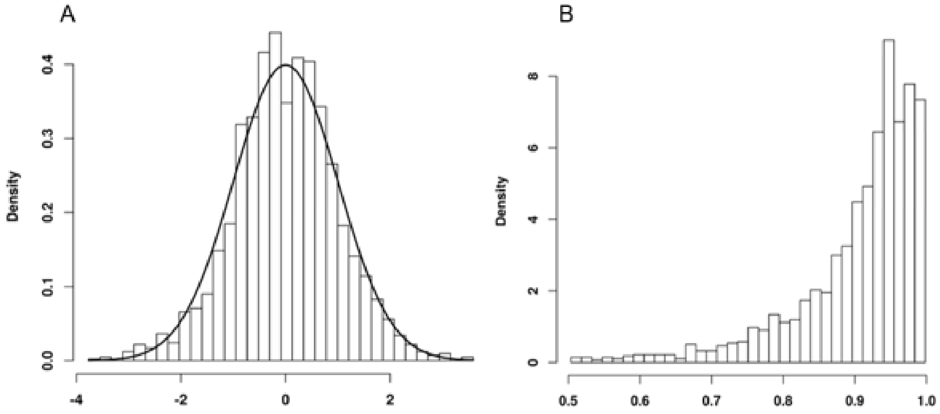


Figure 2-5. Distribution of standardized average daily EBV (A) and its reliability (B) for milk urea concentration of the selected bulls (n = 1,900)

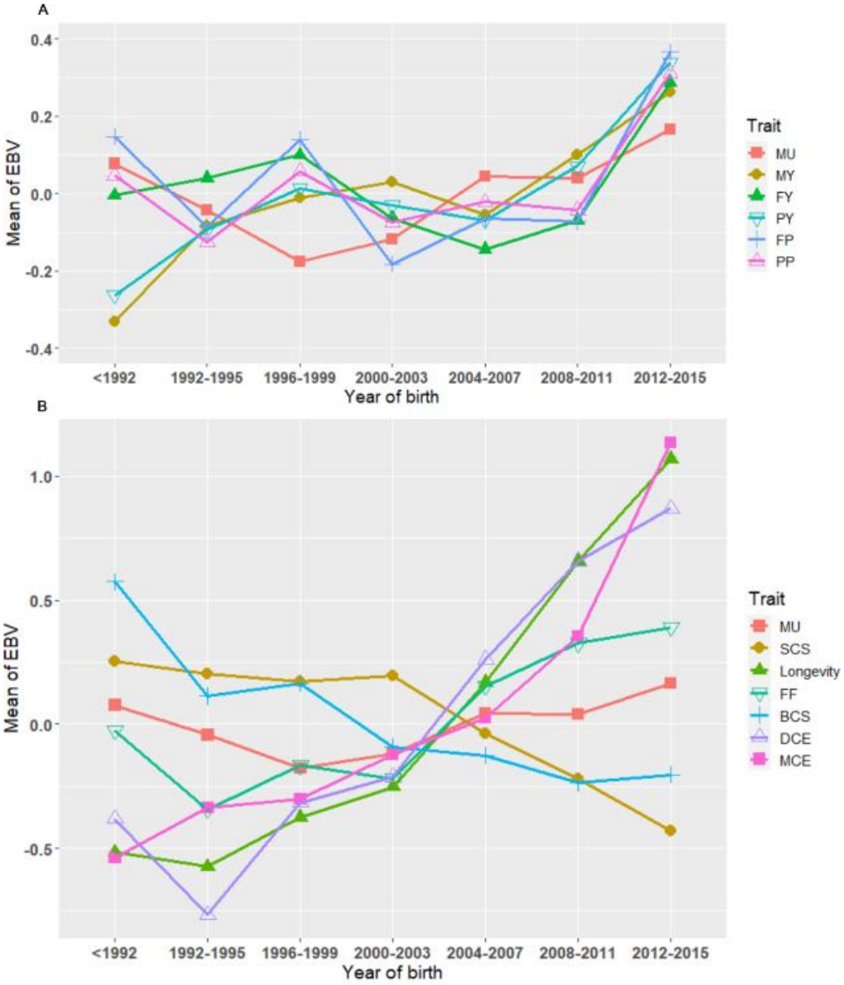


Figure 2-6. Genetic trend of milk urea concentration (MU), milk yield (MY), fat yield (FY), protein yield (PY), SCS, fat percentage (FP), protein percentage (PP) in figure A, and longevity, female fertility (FF), BCS, direct calving ease (DCE), maternal calving ease (MCE) in figure B by year of birth of selected bulls (n = 1,900 for MU, FY, PY, SCS, and n = 1,317 for FP, PP, longevity, FF, BCS, DCE, MCE)

Figure 2-6 shows the genetic trends for the MU, production, longevity, and functional traits involved in this study. The genetic trend of MU gradually decreased from the first to the third group (1996–1999), which reached its minimum level, then gradually increased to the last group (2012–2015) to reach the maximum level. It has been also shown that the N use efficiency of the local cattle population decreased gradually in the past 20 yr. The observed rapid increase in longevity, DCE, and MCE and their positive genetic correlations with MU (the results shown in the next section), may explain a part of this finding. Only SCS showed a declined trend during the last 20 yr; however, indicating that genetic selection has also improved the udder health of the studied cattle population. Moreover, decreasing SCS can be associated with a lower rate of mastitis and therefore higher longevity.

3.5. Approximate Genetic Correlations with Other Traits of Interest

The approximate genetic correlations between MU and MY, FY, PY, and SCS were estimated based on the daily EBV of the traits. The mean (SD) of the REL (5 to 365 DIM) of daily EBV of 1,900 selected bulls for the examined traits ranged from 0.91 (0.04) to 0.95 (0.07). The mean (SD) of the REL of EBV of 1,317 bulls selected for the rest of the traits included (FP, PP, longevity, BCS, FF, DCE, MCE) ranged from 0.72 (0.05) to 0.92 (0.09).

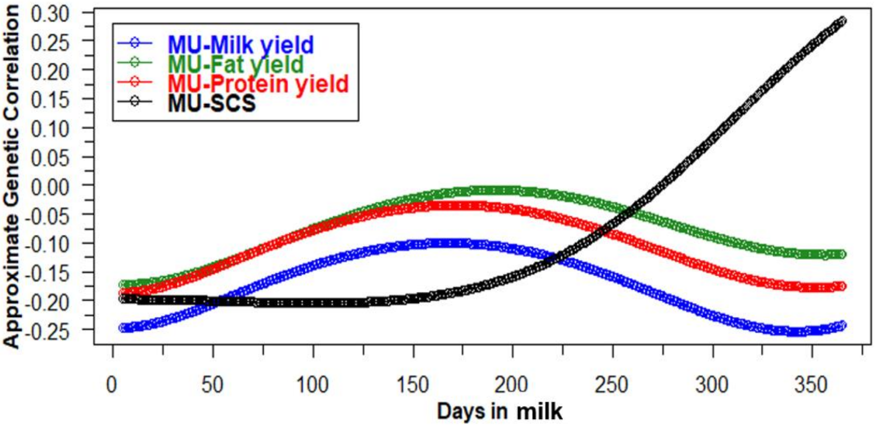


Figure 2-7. Approximate genetic correlation estimated between milk urea concentration (MU), milk yield, fat yield, protein yield, and SCS evaluated by EBV of selected bulls (n = 1,900) over DIM

As shown in Figure 2-7, the range of approximate genetic correlations between MU and production traits (MY, FY, and PY) was found to be from −0.25 to −0.01, indicating that selection for decreasing MU is also associated with increasing milk production traits. Low negative genetic correlation was found between MU and MY, which is in agreement with Samoré et al. (2007). However, Lopez-Villalobos et al. (2018) and Ariyarathne et al. (2021) reported that MU had a moderately positive genetic correlation with milk yield in dairy cows in New Zealand. Only grass-based diets are very common in New Zealand, but not in Belgium, where complementing grazing by energy-rich feedstuff is rather standard. These important differences may cause genetic × environmental interaction that could explain the observed differences. The differences found among the results reported in different studies may be attributed to the difference in the genetic background of the populations studied, and different procedures used. The trait MU was weakly correlated with FY and PY (−0.19 to −0.01). Mucha and Strandberg (2011) reported weak negative to the weak positive genetic correlation between MUN and PY and FY. However, there are studies (e.g., Wood et al., 2003; Stoop et al., 2007) that reported positive genetic correlation between MU and FY or PY. The approximate genetic correlation found between MU and SCS ranged from −0.20 to 0.28 (Figure 2-7). Rzewuska and Strabel (2013) reported a negative genetic correlation between MU and SCS in Polish dairy cattle. Miglior et al. (2007) also found a negative genetic correlation between MUN and SCS for the first 3 parities of Canadian Holstein cows. It is well documented that the N use efficiency in dairy cows decreased (increased MU), but the probability of suffering from mastitis increased (increased SCS) in the late lactation. We speculate that this could be one of the reasons why the genetic correlation between MU and SCS increased from a negative genetic correlation at the beginning of the lactation to a positive genetic correlation in the late part of the lactation.

Table 2-4. Approximate genetic correlation[1] and associated SE between milk urea concentration and selected functional traits evaluated in the selected bulls (n = 1,317)

Trait	DIM 25[2]	DIM 125	DIM 225	DIM 325
Milk composition				
Fat Percentage	0.23 ± 0.03	0.14 ± 0.03	0.15 ± 0.03	0.22 ± 0.03
Protein Percentage	0.20 ± 0.03	0.15 ± 0.03	0.19 ± 0.03	0.26 ± 0.03
Longevity and other functional traits				
Longevity	0.27 ± 0.03	0.27 ± 0.03	0.18 ± 0.03	–0.02 ± 0.03
Female Fertility	0.15 ± 0.03	0.12 ± 0.03	0.09 ± 0.03	0.05 ± 0.03
Body condition score	–0.03 ± 0.03	–0.08 ± 0.03	–0.01 ± 0.04	0.18 ± 0.04
Direct Calving Ease	0.20 ± 0.03	0.25 ± 0.03	0.17 ± 0.03	–0.05 ± 0.03
Maternal Calving Ease	0.21 ± 0.03	0.23 ± 0.03	0.14 ± 0.03	–0.08 ± 0.03

[1] approximate genetic correlation was estimated using the procedure proposed by Blanchard et al. (1983).

[2] EBV of milk urea concentration in corresponding days in milk: 25, 125, 225, or 325

The approximated genetic correlations found between MU and traits including FP, PP, longevity, FF, BCS, DCE, and MCE are presented in Table 2-4. Moreover, we found that the genetic correlations between MU and FP, and PP slightly changed during the lactation. The MU had a positive approximated genetic correlation with FP and PP, which is in agreement with Miglior et al. (2007). However, Bobbo et al. (2020) reported that genetic correlations of MUN and FP or PP were negative, which may be due to the different genetic backgrounds of the breeds considered in their study (i.e., Brown Swiss cows).

The approximated genetic correlation between MU and longevity ranged from −0.02 to 0.27, which is in agreement with Miglior et al. (2006), which used phenotype data to estimate the correlations. Miglior et al. (2006) using 1,568,952 MUN records, reported that an increased MUN was associated with decreased risk of culling in Holstein cows. The approximated genetic correlation found between MU and FF ranged from 0.05 to 0.15, which was consistent with Mucha and Strandberg (2011) and Rzewuska and Strabel (2015). Mucha and Strandberg (2011) used 6, and Rzewuska and Strabel (2015) used 5 indicators for measuring FF, whereas in our study, FF was defined as pregnancy rate (Vanderick et al., 2020). The slightly positive genetic relationship found between MU and FF may be due to the slightly negative correlation between MU and MY and the relatively strong expectation that there is a negative correlation between FF and MY. Siatka et al. (2020) reported that MU had a significant negative effect on the fertility of high-yielding dairy cattle. Moreover, the approximate genetic correlation found between MU and FF was lower at the end of the lactation (325 DIM). In summary, it turns out that MU, longevity, and functional traits (except for BCS) were positively genetically related in the early stage of lactation, and gradually weaken in the later stage.

The approximate genetic correlation found between MU and BCS was close to zero in early and mid-lactation, but a stronger correlation was found in late lactation. Loker et al. (2012) reported that the genetic correlation between MUN and BCS was relatively high in the early and late stages of lactation. High levels of MU in late lactation may indicate excessive protein in the feed, which may improve the BCS of cows. A weakly positive genetic correlation was found between MU and calving ease (direct and maternal) in early lactation.

This research leads to several issues to be considered in future research. First, more than 95% of the cows in this study were Holstein. Strictly speaking, because the phenotypic means of MU and its genetic parameters were reported to be different for specific cattle breeds (e.g., Brown Swiss; Doska et al., 2012; Bobbo et al., 2020), our results can only be considered representative of the Holstein breed in the Walloon Region of Belgium. Therefore, complementary research for dual-purpose Belgian Blue, the second most common breed milked in dairy farms in the Walloon Region, is ongoing simultaneously. Second, genomic selection is now commonly used in dairy cattle, and this is also the case in the Walloon Region. Therefore, we will extend this research toward the use of genomic tools. Finally, this study did not solve the question about the link between MU and N use efficiency. As reported by Chen et al. (2021), we have defined novel predicted N use efficiency traits based on mid-infrared spectral data. Because of the availability of a historical spectral database in the Walloon Region, we are investigating the link between MU and these novel traits. These results will also be useful to develop strategies on how MU EBV should be defined and how they should be used to be useful for reducing environmental impact and improving N efficiency.

4. Conclusions

The genetic parameters of MU were estimated using a random regression model for the first 3 parities. The average daily heritability of MU from DIM 5 to 365 in the first 3 parities ranged from 0.19 to 0.22. High genetic correlations were found among MU in the first 3 parities, so we recommend combining the 3 traits into 1. The approximated genetic correlation found between MU and SCS ranged from −0.20 to 0.28, and the corresponding genetic correlation found between MU and production traits (MY, FY, and PY) ranged from −0.25 to −0.01, indicating that selection for increasing milk production traits and decreasing SCS should lead to decreasing MU. Observation of genetic trends revealed more complex relationships. Genetic correlations between MU and longevity, FF, DCE, and MCE were weakly positive in early lactation (0.12–0.27). The findings of this study can be used as the first step in developing a routine genetic evaluation for MU and its inclusion into the genetic selection program in the Walloon Region of Belgium.

5. Acknowledgments

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Yansen Chen. The authors acknowledge the support of the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE; Namur, Belgium) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech (Gembloux, Belgium) provided by the technical plat-form Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grants No. T.0095.19 (PDR “DEEPSELECT”) and J.0174.18 (CDR “PREDICT-2”). Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI) funded by the National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grant No. 2.5020.11. The authors acknowledged the support of the Walloon Breeder Association (awé group, Belgium). The authors have not stated any conflicts of interest.

6. References

Ariyarathne, H.B.P.C., M. Correa-Luna, H.T. Blair, D.J. Garrick, and N. Lopez-

Villalobos. 2019. Estimation of heritabilities for milk urea and efficiency of crude protein utilization by day of lactation in a New Zealand dairy cow herd. New Zeal. J. Anim. Sci. Prod. 79:183–187.

Ariyarathne, H.B.P.C., M. Correa-Luna, H.T. Blair, D.J. Garrick, and N. Lopez-Villalobos. 2021. Genetic parameters for efficiency of crude protein utilisation and its relationship with production traits across lactations in grazing dairy cows. New Zeal. J. Agric. Res. 64:62–82. https://doi.org/10.1080/00288233.2020.1726414.

Bastin, C., L. Laloux, A. Gillon, F. Miglior, H. Soyeurt, H. Hammami, C. Bertozzi, and N. Gengler. 2009. Modeling milk urea of Walloon dairy cows in management perspectives. J. Dairy Sci. 92:3529–3540. https://doi.org/10.3168/jds.2008-1904.

Beatson, P.R., S. Meier, N.G. Cullen, and H. Eding. 2019. Genetic variation in milk urea nitrogen concentration of dairy cattle and its implications for reducing urinary nitrogen excretion. Animal 13:2164–2171. https://doi.org/10.1017/S1751731119000235.

Bendelja, D., Z. Prpić, N. Mikulec, Z. Ivkić, J. Havranek, and N. Antunac. 2011. Milk urea concentration in Holstein and Simmental cows. Mljekarstvo 61:45–55. https://hrcak.srce.hr/65204.

Blanchard, P.J., R.W. Everett, and S.R. Searle. 1983. Estimation of genetic trends and correlations for Jersey cattle. J. Dairy Sci. 66:1947–1954. https://doi.org/10.3168/jds.S0022-0302(83)82033-5.

Bobbo, T., C. Roveglia, M. Penasa, G. Visentin, R. Finocchiaro, and M. Cassandro. 2019. Genetic relationships of alternative somatic cell count traits with milk yield, composition and udder type traits in Italian Jersey cows. Anim. Sci. J. 90:808–817. https://doi.org/10.1111/asj.13204.

Bobbo, T., M. Penasa, A. Rossoni, and M. Cassandro. 2020. Short communication: Genetic aspects of milk urea nitrogen and new indicators of nitrogen efficiency in dairy cows. J. Dairy Sci. 103:9207–9212. https://doi.org/10.3168/jds.2020-18445.

Calo, L.L., R.E. McDowell, L.D. VanVleck, and P.D. Miller. 1973. Genetic Aspects of Beef Production among Holstein-Friesians pedigree selected for milk production. J. Anim. Sci. 37:676–682. https://doi.org/10.2527/jas1973.373676x.

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.

Čobanović, K., D. Kučević, M. Plavšić, and V. Bogdanović. 2017. Impact of non nutritional factors on milk urea concentration and its relationship with production and fertility traits in Vojvodina dairy herds. Mljekarstvo 67:267–276. https://doi.org/10.15567/mljekarstvo.2017.0404.

Correa-Luna, M., D. Donaghy, P. Kemp, M. Schutz, and N. Lopez-villalobos. 2019. Effect of genetic merit for milk urea on milk production and efficiency of crude protein utilization of grazing cows with contrasting supplement inclusion. New Zeal. J. Anim. Sci. Prod. 79:112–117.

Croquet, C., P. Mayeres, A. Gillon, S. Vanderick, and N. Gengler. 2006. Inbreeding depression for global and partial economic indexes, production, type, and functional traits. J. Dairy Sci. 89:2257–2267. https://doi.org/10.3168/jds.S0022-0302(06)72297-4.

Doska, M.C., D.F.F. Silva, J.A. Horst, A.A. Valloto, P. Rossi, and R. de Almeida. 2012. Sources of variation in milk urea nitrogen in Paraná dairy cows. R. Bras. Zootec. 41:692–697. http://dx.doi.org/10.1590/S1516-35982012000300032.

Fatehi, F., A. Zali, M. Honarvar, M. Dehghan-banadaky, A.J. Young, M. Ghiasvand, and M. Eftekhari. 2012. Review of the relationship between milk urea nitrogen and days in milk, parity, and monthly temperature mean in Iranian Holstein cows. J. Dairy Sci. 95:5156–5163. https://doi.org/10.3168/jds.2011-4349.

Gengler, N., A. Tijani, G.R. Wiggans, and I. Misztal. 1999. Estimation of (Co)variance function coefficients for test day yield with a expectation-maximization restricted maximum likelihood algorithm. J. Dairy Sci. 82:1849.e1-1849.e23. https://doi.org/10.3168/jds.S0022-0302(99)75417-2.

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, and 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.

Glatz-Hoppe, J., A. Boldt, H. Spiekers, E. Mohr, and B. Losand. 2020. Relationship between milk constituents from milk testing and health, feeding, and metabolic data of dairy cows. J. Dairy Sci. 103:10175–10194. https://doi.org/10.3168/jds.2019-17981.

Hojman, D., M. Gips, and E. Ezra. 2005. Association between live body weight and milk urea concentration in Holstein cows. J. Dairy Sci. 88:580–584. https://doi.org/10.3168/jds.S0022-0302(05)72721-1.

Hossein-Zadeh, N.G., and M. Ardalan. 2011. Genetic relationship between milk urea nitrogen and reproductive performance in Holstein dairy cows. Animal 5:26–32. https://doi.org/10.1017/S1751731110001606.

Hristov, A.N., A. Bannink, L.A. Crompton, P. Huhtanen, M. Kreuzer, M. McGee, P. Nozière, C.K. Reynolds, A.R. Bayat, D.R. Yáñez-Ruiz, J. Dijkstra, E. Kebreab, A. Schwarm, K.J. Shingfield, and Z. Yu. 2019. Invited review: Nitrogen in ruminant nutrition: A review of measurement techniques. J. Dairy Sci. 102:5811–5852. https://doi.org/10.3168/jds.2018-15829.

Jonker, J.S., R.A. Kohn, and R.A. Erdman. 1998. Using Milk Urea Nitrogen to Predict Nitrogen Excretion and Utilization Efficiency in Lactating Dairy Cows. J. Dairy Sci. 81:2681–2692. https://doi.org/10.3168/jds.S0022-0302(98)75825-4.

Kananub, S., P. Pechkerd, J. VanLeeuwen, H. Stryhn, and P. Arunvipas. 2020. Evaluation of influence of milk urea nitrogen on reproductive performance in smallholder dairy farms. Aust. Vet. J. 98:375–379.

https://doi.org/10.1111/avj.12946.

Kauffman, A.J., and N.R. St-Pierre. 2001. The relationship of milk urea nitrogen to urine nitrogen excretion in Holstein and Jersey cows. J. Dairy Sci. 84:2284–2294. https://doi.org/10.3168/jds.S0022-0302(01)74675-9.

König, S., Y.M. Chang, U.U.V. Borstel, D. Gianola, and H. Simianer. 2008. Genetic and phenotypic relationships among milk urea nitrogen, fertility, and milk yield in Holstein cows. J. Dairy Sci. 91:4372–4382. https://doi.org/10.3168/jds.2008-1236.

Loker, S., C. Bastin, F. Miglior, A. Sewalem, L.R. Schaeffer, J. Jamrozik, A. Ali, and V. Osborne. 2012. Genetic and environmental relationships between body condition score and milk production traits in Canadian Holsteins. J. Dairy Sci. 95:410–419. https://doi.org/10.3168/jds.2011-4497.

Lopez-villalobos, N., M. CorreaLuna, J. Burke, N. Sneddon, M. Schultz, D. Donaghy, and P. Kemp. 2018. Genetic parameters for milk urea concentration and milk traits in New Zealand grazing dairy cattle. New Zeal. J. Anim. Sci. Prod. 78:56–61.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2020. Grazing dairy cows with low milk urea nitrogen breeding values excrete less urinary urea nitrogen. Sci. Total Environ. 739:1-8. https://doi.org/10.1016/j.scitotenv.2020.139994.

Mayeres, P., J. Stoll, J. Bormann, R. Reents, and N. Gengler. 2004. Prediction of daily milk, fat, and protein production by a random regression test-day model. J. Dairy Sci. 87:1925–1933. https://doi.org/10.3168/jds.S0022-0302(04)73351-2.

Miglior, F., A. Sewalem, J. Jamrozik, D.M. Lefebvre, and R.K. Moore. 2006. Analysis of milk urea nitrogen and lactose and their effect on longevity in Canadian dairy cattle. J. Dairy Sci. 89:4886–4894. https://doi.org/10.3168/jds.S0022-0302(06)72537-1.

Miglior, F., A. Sewalem, J. Jamrozik, J. Bohmanova, D.M. Lefebvre, and R.K. Moore. 2007. Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in Canadian Holstein cattle. J. Dairy Sci. 90:2468–2479. https://doi.org/10.3168/jds.2006-487.

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 Jan. 06, 2021. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Mitchell, R.G., G.W. Rogers, C.D. Dechow, J.E. Vallimont, J.B. Cooper, U. Sander-Nielsen, and J.S. Clay. 2005. Milk urea nitrogen concentration: Heritability and genetic correlations with reproductive performance and disease. J. Dairy Sci. 88:4434–4440. https://doi.org/10.3168/jds.S0022-0302(05)73130-1.

Mucha, S., and E. Strandberg. 2011. Genetic analysis of milk urea nitrogen and relationships with yield and fertility across lactation. J. Dairy Sci. 94:5665–5672. https://doi.org/10.3168/jds.2010-3916.

Roy, B., R.K. Mehla, and S.K. Sirohi. 2003. Influence of milk yield, parity, stage of lactation and body weight on urea and protein concentration in milk of Murrah buffaloes. Asian-Australas. J. Anim. Sci. 16:1285–1290. https://doi.org/10.5713/ajas.2003.1285.

Rzewuska, K., and T. Strabel. 2013. Genetic parameters for milk urea concentration and milk traits in Polish Holstein-Friesian cows. J. Appl. Genet. 54:473–482. https://doi.org/10.1007/s13353-013-0159-8.

Rzewuska, K., and T. Strabel. 2015. The genetic relationship between reproduction traits and milk urea concentration. Anim. Sci. Pap. Reports 33:243–256.

Samoré, A.B., C. Romani, A. Rossoni, E. Frigo, O. Pedron, and A. Bagnato. 2007. Genetic parameters for casein and urea content in the Italian Brown Swiss dairy cattle. Ital. J. Anim. Sci. 6:201–203. https://doi.org/10.4081/ijas.2007.1s.201.

Satoła, A., E. Ptak, A. Otwinowska-Mindur, and W. Jagusiak. 2017a. Genetic parameters for lactose percentage and urea concentration in milk of polish Holstein–Friesian cows. Anim. Sci. Pap. Reports 35:159–172.

Satoła, A., E. Ptak, W. Jagusiak, and A. Otwinowska-Mindur. 2017b. Genetic relationship of conformation traits with lactose percentage and urea concentration in milk of Polish Holstein-Friesian cows. Anim. Sci. Pap. Reports 35:241–252.

Siatka, K., A. Sawa, M. Bogucki, and S. Krezel-Czopek. 2020. The effect of milk urea level on fertility parameters in Holstein–Friesian dairy cows. Turkish J. Vet. Anim. Sci. 44:42–46. https://doi.org/10.3906/vet-1810-41.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, and A. Bannink. 2013. A review of factors influencing milk urea concentration and its relationship with urinary urea excretion in lactating dairy cattle. J. Agric. Sci. 151:407–423. https://doi.org/10.1017/S0021859612000561.

Stoop, W.M., H. Bovenhuis, and J.A.M. Van Arendonk. 2007. Genetic parameters for milk urea nitrogen in relation to milk production traits. J. Dairy Sci. 90:1981–1986. https://doi.org/10.3168/jds.2006-434.

Stranden, I., M. Lidauer, E. A. Mäntysaari, and J. Pösö. 2000. Calculation of Interbull weighting factors for the Finnish test day model. Interbull Bull. 26:78–79.

Vanderick S., R. R. Mota, K. Wijnrocs, and N. Gengler. 2020. Description of the genetic evaluation systems used in the Walloon Region of Belgium. Accessed Jan. 06, 2021. http://www.elinfo.be/docs/GESen2012.pdf.

Wood, G.M., P.J. Boettcher, J. Jamrozik, G.B. Jansen, and D.F. Kelton. 2003. Estimation of genetic parameters for concentrations of milk urea nitrogen. J. Dairy Sci. 86:2462–2469. https://doi.org/10.3168/jds.S0022-0302(03)73840-5.

Yin, T., B. Bapst, U.U.V. Borstel, H. Simianer, and S. König. 2012. Genetic parameters for gaussian and categorical traits in organic and low input dairy cattle herds based on random regression methodology. Livest. Sci. 147:159–169. https://doi.org/10.1016/j.livsci.2012.04.017.

3

Chapter III Estimation of genetic parameters for predicted nitrogen use efficiency and losses in early lactation of Holstein cows

Adapted from: Chen, Y., S. Vanderick, R.R. Mota, C. Grelet, G. 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. 104:4412-4423. https://doi.org/10.3168/jds.2020-18849.

Foreword

While the previous Chapter discussed the traditional proxy (milk urea concentration) for nitrogen use efficiency (NUE), this Chapter presented a genetic analysis of two novel proxies - predicted NUE (PUE) and nitrogen losses (PNL). The PNUE and PNL are directly related to NUE and nitrogen pollution, respectively, and may be more representative of the real situation of dairy cows. These results provide the first references for PNUE and PNL to inform breeding and management strategies to improve farmers' profits and reduce environmental pollution.

Abstract

The objective of this study was to estimate genetic parameters of predicted N use efficiency (PNUE) and N losses (PNL) as proxies of N use and loss for Holstein cows. Furthermore, we have assessed approximate genetic correlations between PNUE, PNL, and dairy production, health, longevity, and conformation traits. These traits are considered important in many countries and are currently evaluated by the International Bull Evaluation Service (Interbull). The values of PNUE and PNL were obtained by using the combined milk mid-infrared (MIR) spectrum, parity, and milk yield–based prediction equations on test-day MIR records with days in milk (DIM) between 5 and 50 d. After editing, the final data set comprised 46,163 records of 21,462 cows from 154 farms in 5 countries. Each trait was divided into primiparous and multiparous (including second to fifth parity) groups. Genetic parameters and breeding values were estimated by using a multi-trait (2-trait, 2-parity classes) repeatability model. Herd-year-season of calving, DIM, age of calving, and parity were used as fixed effects. Parity (within-parity permanent environment), nongenetic cow (across-parity permanent environment), additive genetic animal, and residual effects were defined as random effects. The estimated heritabilities of PNUE and PNL ranged from 0.12 to 0.14, and the repeatabilities ranged from 0.40 to 0.55, respectively. The estimated genetic correlations between PNUE and PNL were negative and high (from −0.89 to −0.53), whereas the phenotypic correlations were also negative but relatively low (from −0.45 to −0.11). At a level of reliability of more than 0.30 for all novel traits, a total of 504 bulls born after 1995 had also publishable Interbull multiple-trait across-country estimated breeding values (EBV). The approximate genetic correlations between PNUE and the other 30 traits of interest, estimated as corrected correlations between EBV of bulls, ranged from −0.46 (udder depth) to 0.47 (milk yield). Obtained results showed the complex genetic relationship between efficiency, production, and other traits: for instance, more efficient cows seem to give more milk, which is linked to deeper udders, but seem to have lower health, fertility, and longevity. Additionally, the approximate genetic correlations between PNL, lower values representing less loss of N, and the 30 other traits, were from −0.32 (angularity) to 0.57 (direct calving ease). Even if further research is needed, our results provided preliminary evidence that the PNUE and PNL traits used as proxies could be included in genetic improvement programs in Holstein cows and could help their management.

Key words: genetic correlation, heritability, indirect selection, repeatability model, mid-infrared spectrum

1. Introduction

The concept of efficiency of dairy cows has to be put in a holistic perspective, as efficiency might enter a tradeoff with the health status and the ability of the cow to reproduce (Figure 3-1). Therefore, efficiency is also linked to the well-being, robustness, and longevity of the cows. Moreover, efficiency is not only feed efficiency (Hayes and Ageeb, 2002) but also nitrogen use efficiency (NUE), which is considered one of the most important indicators of ruminant feed utilization (Wheadon et al., 2014). The protein in the feed is the most expensive raw material, so the increase in NUE is beneficial to the interests of farmers (Cantalapiedra-Hijar et al., 2018). Nitrogen emissions from the dairy industry have been identified as important factors causing groundwater and surface water pollution, and greenhouse gas emissions (Adenuga et al., 2019). The reduction of N emissions would protect the environment, and we might expect that environmental losses are also linked to efficiency (Figure 3-1).

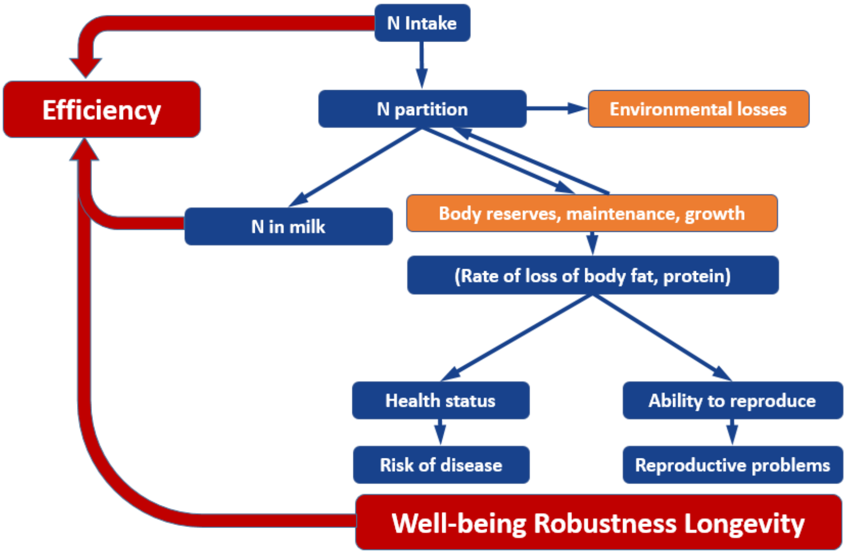


Figure 3-1. Holistic view of the concept of efficiency in the context of N (in red) and relationship between N in milk and N intake (N loss shown in orange boxes) in dairy cows. (Modified from Gengler et al., 2018.)

In this research, we studied 2 traits: NUE, defined as the ratio of grams of N in milk per grams of N intake, and nitrogen loss (NL), defined as grams of N intake minus grams of N in milk (Phuong et al., 2013; Grelet et al., 2020). It has to be noted that the concept of NL used hereafter is a simple difference between N intake and N in milk. Therefore, it does only reflect imperfect environmental losses (Figure 3-1). Concerning NUE, ratio traits are not preferred in genetics because of 2 issues: a potentially problematic, non-normal, distribution of records, and the difficulty of attributing changes to the numerator (here output grams of N in milk) or the denominator (here input grams of N intake). However, we studied both traits because NUE is often preferred for management purposes. As already stated, the improvement of the utilization of N in cattle may have an adverse effect on their health and longevity because it affects also N available for body reserves, maintenance needs, and even growth for young animals (Figure 3-1). Therefore, we also tried to study the genetic relationships between N efficiency traits and other important traits such as milk yields and composition, health, longevity, other functional traits, and conformation traits.

Literature has reported that NUE shows large variability between individuals and herds, with values between 8% and 42% (Castillo et al., 2001). This suggests potential opportunities to improve NUE under the condition of sufficient genetic variance, by genetic selection. To our knowledge, genetic parameters and the correlation between NUE and NL, and with other traits in Holstein cows, have not yet been analyzed. However, the efficiency of crude protein utilization (ECPU, CP in milk divided by CP intake), as reported by Lopez-Villalobos et al. (2018), was a trait that can be considered similar to NUE (ratio of grams of N in milk derived from CP in milk and grams of N intake derived from CP intake). Alternatively, ^{15}N has been used to study the NUE and NL of cows (Cantalapiedra-Hijar et al., 2016, 2018; Nasrollahi et al., 2019). Because of its high cost and complicated operation, it is impossible to collect the large amount of data needed for genetic evaluation or even genetic research. Therefore, obtaining large numbers of direct NUE and NL, as well as ECPU records, is difficult, and the largest study, by Lopez-Villalobos et al. (2018), estimated the heritability (h^2) of ECPU, using only 2,896 records. These facts highlight the urgent need for large-scale and easy-to-measure proxies for N (or protein) efficiency traits.

Recently, Grelet et al. (2020) established models to obtain proxies for NUE (PNUE) and NL (PNL) based on their prediction by mid-infrared (MIR) spectra. The use of spectra from the routine MIR analysis of milk is fast and simple, and allows traits to be predicted on a large scale. Predictions can be redone whenever new or better equations become available (Gengler et al., 2016). As emphasized by many recent papers (e.g., Vanlierde et al., 2015; Gengler et al., 2016), the goal of using MIR is to establish predictors that can be used as proxies for the unavailable direct traits, not the perfect replacement of a direct trait. For our traits, the correlations of 0.82 for PNUE and 0.77 for PNL observed by Grelet et al. (2020) in their validations, and defined as the square root of these validation coefficients of determination (R^2), can be considered as being approximations of the level of phenotypic correlations expected between direct and our proxy traits.

The equations developed by Grelet et al. (2020) to predict PNUE and PNL were applied to a much larger data set. The objectives of this study were (1) to estimate genetic parameters of PNUE and PNL for Holstein cows in this data set and (2) to assess approximate genetic correlations between PNUE, PNL, and important traits, as milk yield and composition, health, longevity, and other functional traits, and conformation traits (30 traits).

2. Materials and methods

2.1. Data Collection and Editing

The data used in this study were collected by the Genotype Plus Environment (GplusE) FP7-Project (http://www.gpluse.eu). The initial data set, including milk MIR, parity, and milk yield, comprised a total of 52,065 records from 154 herds from 5 partner countries of the GplusE project. This specific population was also studied because the cows used in this study were genotyped during GplusE, a fact that will be important in genomic follow-up studies.

The original data were edited to select records between 5 and 50 DIM, and observed during the first 5 parities. This allows us to keep the data similar to those used for the reference population. All milk MIR spectra were standardized according to Grelet et al. (2015). The PNUE and PNL of each cow were predicted by the equations based on the models established by Grelet et al. (2020), with milk MIR spectra, milk yield, and parity as additional predictors, using support vector machine regression. The determination coefficients and root mean square errors of validation were 0.68 and 5.01% for the PNUE model, and 0.59 and 0.07 kg/d for the PNL model, respectively.

Allowed minimum and maximum age depended on parities. For parity 1 the range was from 22 to 36, for parity 2 from 34 to 50, for parity 3 from 46 to 64, for parity 4 from 58 to 78, and for parity 5 from 70 to 92 mo of age at calving were included. Finally, the final data set contained 46,163 records of 21,462 cows from 154 farms. The distribution of the final data across 5 organizations and universities was as follows: Royal Veterinary College (London, UK; n = 666), Walloon Breeders Association (Ciney, Belgium; n = 20,144), Irish Cattle Breeding Federation Society Limited (Bandon, Ireland; n = 3,543), Huazhong Agricultural University (Wuhan, China; n = 188), and CRV (Arnhem, the Netherlands; n = 21,622). The records were divided into 2 parity classes according to lactation numbers 1 and 2 through 5. Hereafter, these 4 traits are identified as PNUE1, PNL1, PNUE2+, and PNL2+. The pedigree related to the data sets comprised 75,700 animals.

As stated by Grelet et al. (2020), the use of MIR equations in other populations than the ones used for calibration is dangerous. Because of the lack of feeding data, we could not directly validate the predictions obtained using the equations in our population. However, by applying 2 filters, one on the MIR data and one on the values of the predictions, we acted very prudently. First, the population where the prediction equations are used should also be spectrally similar to the calibration population. A widely used strategy is to compute the global H parameter, which is based on the standardized Mahalanobis distance between the MIR record to be used and the calibration data sets (Whitfield et al., 1987). The global H is therefore a multidimensional generalized measure of how many standard deviations (SD) away a MIR record is from the mean of the corresponding calibration population. By using only spectra with a global $H \leq 3$, we have limited the use of the equations to MIR records close to the calibration population, and therefore we have minimized the risk of using equations improperly. In addition, the means and SD of the values predicted for the 4 traits PNUE1, PNL1, PNUE2+, and PNL2+ were established, and records were kept only when they were inside a ≤ 3 SD range. Even if these 2 measures cannot guarantee the applicability of the equations, these measures not only established that our spectra were very close to the calibration spectra but also that the predicted values stayed in a range compatible with the trait definitions of Grelet et al. (2020)—that is, inside the range of the values used for calibration.

2.2. Model

A multitrait (4 traits; 2 traits, 2 parity classes) repeatability model was used for estimating variance components and breeding values. The model was fitted as follows:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Xb} + \mathbf{Qq} + \mathbf{W_1c} + \mathbf{W_2p} + \mathbf{Za} + \mathbf{e},$$

where **y** was the vector of 4 traits PNUE1, PNL1, PNUE2+, and PNL2+. For each trait, **h** was the vector of fixed herd-year-season of calving classes (4 seasons from December to February, March to May, June to August, and September to November); **b** was the vector of fixed regression coefficients for DIM, after standardization, and its quadratic; **q** was the vector of fixed regression coefficients of the age of calving, after standardization, defined as a constant (parity effect), linear, and quadratic regression, defined within parities (1–5); **c** was a vector of the nongenetic cow (across-parity permanent environment) random effects; **p** was a vector of nongenetic cow × parity (within-parity permanent environment) random effects, modeled only for PNUE2+ and PNL2+, as they allowed us to distinguish records for the same cow occurring during different parities (second to fifth parity); **a** was a vector of random additive genetic effects; and **e** was a vector of random residual effects. Additionally, **H**, **X**, **Q**, $\mathbf{W_1}$, $\mathbf{W_2}$, and **Z** were incidence matrices assigning observations to effects.

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

$$\mathrm{E}(\mathbf{y}) = \mathbf{Hh} + \mathbf{Xb} + \mathbf{Qq},$$

$$\mathrm{E}(\mathbf{c}) = \mathrm{E}(\mathbf{p}) = \mathrm{E}(\mathbf{a}) = \mathrm{E}(\mathbf{e}) = 0,$$

The matrices $\mathrm{V}(\mathbf{c})$ and $\mathrm{V}(\mathbf{a})$ contained blocks of 4 × 4 elementary (co)variances between the 4 traits. For $\mathrm{V}(\mathbf{e})$, off-diagonals were modeled as being zero, these elements representing residual covariances between the respective traits. By this approach, environmental covariances between traits were concentrated in permanent environmental effects. For $\mathrm{V}(\mathbf{p})$ the elementary (co) variance matrix was reduced to a 2 × 2 matrix, because only the (co)variances associated with PNUE2+ and PNL2+ were present; hereafter, we have to emphasize that this (co)variance does not exist for PNUE1 and PNL1.

2.3. Variance Components, Heritabilities, and Repeatabilities

Computations were performed using the BLUPF90 family of programs (Misztal et al., 2018). Variance components were estimated by using first the REMLF90 (version 1.84) program and then, at convergence, approximate standard errors of all calculated parameters were obtained according to the algorithm of Meyer and Houle (2013) as implemented in the AIREMLF90 (version 1.144) program. For each trait PNUE1 and PNL1, h^2 was defined as $\sigma_a^2/(\sigma_a^2 + \sigma_c^2 + \sigma_e^2)$ and for PNUE2+ and PNL2+, h^2 was defined as $\sigma_a^2/(\sigma_a^2 + \sigma_c^2 + \sigma_p^2 + \sigma_e^2)$ where σ_a^2 was the additive genetic variance; σ_c^2 was the across-parity permanent environment (non-genetic cow) variance; for PNUE2+ and PNL2+, σ_p^2 was the within parity permanent environment (non-genetic cow x parity) variance and σ_e^2 was the residual variance again defined for each trait. Permanent environmental covariance between test-day records of a cow was assumed to be σ_c^2 across parities and $\sigma_c^2 + \sigma_p^2$ inside a given parity for traits PNUE2+ and PNL2+. As traits PNUE1 and PNL1 were defined inside the first parity only, the associated permanent environmental covariance was limited to σ_c^2. For PNUE1 and PNL1, repeatability within the first parity was defined as constant parameters across the range of DIM in early lactation as $(\sigma_a^2 + \sigma_c^2)/(\sigma_a^2 + \sigma_c^2 + \sigma_e^2)$. For PNUE2+ and PNL2+, within parity repeatability was defined as a constant parameter across the range of DIM in early lactation, as $(\sigma_a^2 + \sigma_c^2 + \sigma_p^2)/(\sigma_a^2 + \sigma_c^2 + \sigma_p^2 + \sigma_e^2)$.

2.4. Genetic, Phenotype, and Approximate Genetic Correlations

Genetic correlations were computed directly, using the estimated genetic covariances and variances, as the ratio of the covariance to the square root of the product of the corresponding variances. Phenotype correlations were computed similarly. Required phenotypic covariances and variances were obtained as the sum of relevant (co)variance components.

Breeding values were estimated by using the BLUPF90 (version 1.68) program, and reliabilities were computed from extracted diagonal elements of the inverted coefficient matrix. The approximate genetic correlations between PNUE, PNL, and other traits of interest were estimated using the strategy presented by Blanchard et al. (1983), which is a generalization of Calo's method when groups of bulls had very diverse reliability. We identified bulls born after 1995 (included) in the pedigree, having reliabilities for all 4 novel traits that were above 0.3 and having also publishable International Bull Evaluation Service (Interbull) multiple-trait across-country evaluation–based EBV on Walloon scales (https://www.elinfo.be/telechargerEN.html). Based on a suggestion by A. Legarra (INRAE, Toulouse, France, personal communication) SE of these approximate genetic correlation estimates were estimated using 1,000 bootstrapped replicates. In the bootstrap procedure, we used the number of selected bulls to sample from all selected bulls with replacement.

Additional data preparation and processing were done using R (R Core Team, 2020; version 3.6.3), data. table package (Dowle and Srinivasan, 2019; version 1.12.8), and boot package (Canty and Ripley, 2019; version 1.3-24). The graphs were made using the ggplot2 package (Wickham, 2016; version 3.30).

3. Results and discussion

3.1. Descriptive Statistics

The distribution of PNUE and PNL are shown in Figure 3-2. Under visual inspection, the range of values was acceptably normally distributed. Within the same parity, the variability of PNUE and PNL remained large. Mean values for multiparous cows increased 14.86% and 12.41% over the primiparous Holstein cows, respectively. As explained previously, based on observed means and SD, observations outside 3 SD from the means for each trait were considered as outliers and were not used in the study. For PNUE1, PNL1, PNUE2+, and PNL2+ we eliminated only 40, 161, 69, and 166 records, respectively. This restricted the range of values for PNUE1 and PNUE2+ to between 13.6 and 63.5%, which is inside the values used by Grelet et al. (2020) in the calibration process (individual daily NUE ranging from 9.7 to 81.7%, with an average of 36.9% and SD of 10.4%).

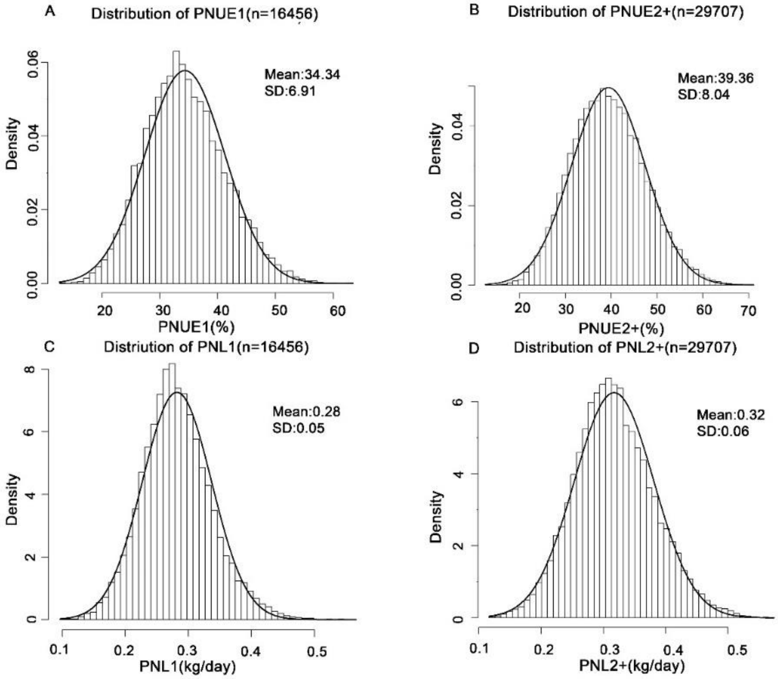


Figure 3-2. Distribution of predicted N use efficiency (A and B) and predicted N losses (C and D) of Holstein cows. PNUE1 = predicted N use efficiency in first-parity Holstein cows; PNL1 = predicted N losses in first-parity Holstein cows; PNUE2+ = predicted N use efficiency in second- to fifth-parity Holstein cows; PNL2+ = predicted N losses in second- to fifth-parity Holstein cows

As shown in Figure 3-3, we found that as the parity (the first 5 parities) increases, the average PNUE gradually increases, and the PNL increases first, before becoming stable and then showing a slight tendency to decrease after parity 3. For PNL within each parity, we found a downward trend with the increase of calving age. All these results suggest that even if cows in higher parities tend to be more efficient (higher PNUE), for PNL later parities compared with the first parity showed higher N losses. Also, although older calving ages inside a given parity were not good for efficiency in the first parity, older animals inside a given parity showed lower losses. These results supported the idea that although the definitions of PNUE and PNL may indicate similarities, they do not describe the same biological background.

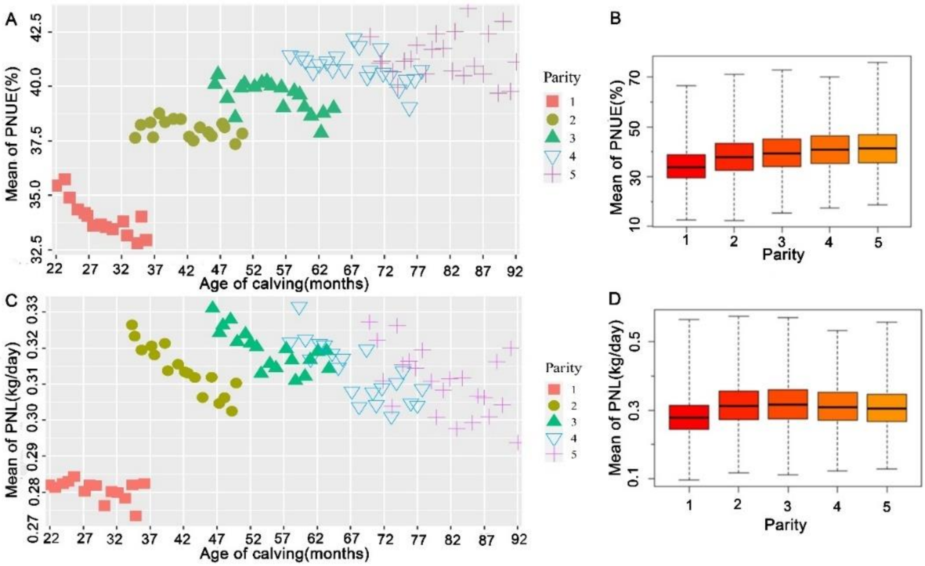


Figure 3-3. Changes in average predicted N use efficiency (PNUE) and predicted N losses (PNL) with calving age (A and C) and parity of Holstein cows (B and D)

3.2. Variance Components, Heritabilities, and Repeatabilities

The variance for each component, h^2, and repeatability of the four traits are shown in Table 3-1. The h^2 of the four traits ranged from 0.115 to 0.144. Given these intermediate levels of h^2, which are similar to h^2 found for SCS, the hypothesis considering PNUE or PNL as traits that could be selected, can be defended (Montaldo et al., 2010). Repeatability ranged between 0.395 and 0.550. It can be seen that the h^2 and repeatability of PNUE and PNL increased slightly from the first to the later parity and that values for PNUE tended to be higher than PNL. Because we are the first to report the genetic parameters of PNUE and PNL, there is no directly comparable literature. However, as there were studies that used ECPU for protein efficiency (Hayes and Ageeb, 2002; Ariyarathne et al., 2019), so we compare them with our results. Recently, Ariyarathne et al. (2019) estimated the average h^2 of ECPU to be 0.16, which was similar to our results (from 0.12 to 0.14), but their mean repeatability tended to be lower (0.38 vs. 0.55).

Table 3-1. Heritability (h^2), repeatability, additive genetic variance (σ^2_a), across-parity permanent environment (nongenetic cow) variance (σ^2_c), within parity permanent environment (non-genetic cow × parity) variance (σ^2_p) (only for second and later lactations), and residual variance (σ^2_e) of the proxies for N use efficiency (%) and losses (expressed as dekagram/day) in primiparous (n = 16,456) and multiparous (n = 29,707) Holstein cows

Trait[1]	h^2	Repeatability	σ^2_a	σ^2_c	σ^2_p	σ^2_e
PNUE1	0.131 ± 0.019	0.487 ± 0.011	3.42 ± 0.52	9.31 ± 0.55	NA[2]	13.39 ± 0.26
PNL1	0.115 ± 0.017	0.395 ± 0.012	1.94 ± 0.30	4.71 ± 0.34	NA	10.17 ± 0.19
PNUE2+	0.144 ± 0.015	0.550 ± 0.007	4.82 ± 0.52	1.80 ± 0.53	11.72 ± 0.43	15.02 ± 0.23
PNL2+	0.129 ± 0.013	0.430 ± 0.009	3.19 ± 0.34	0.68 ± 0.36	6.76 ± 0.33	14.12 ± 0.20

[1] PNUE1 = predicted N use efficiency in first-parity Holstein cows; PNL1 = predicted N losses in first-parity Holstein cows; PNUE2+ = predicted N use efficiency in second- to fifth-parity Holstein cows; PNL2+ = predicted N losses in second- to fifth-parity Holstein cows.

[2] NA = not applicable.

A total of 504 bulls were used in the estimation of approximate genetic correlations. Because Ireland and Belgium are currently importers of a lot of Holstein semen, many important semen-exporting countries were present (e.g., the USA, Canada, Italy, Germany, and France). In the Netherlands, many local bulls are traditionally used, adding to the strong presence of bulls born in this country. The precise distributions, according to the countries of origin and year of birth of the bulls, are shown in Table 3-2. Breeding values used in this study showed averages and SD for the reliability of PNUE1, PNL1, PNUE2+, and PNL2+ that ranged from 0.47 to 0.52 and from 0.14 to 0.16. The averages and SD of reliability for EBV of all multiple-trait across-country evaluation traits ranged from 0.75 to 0.96 and from 0.03 to 0.14, respectively.

Table 3-2. Distribution of year and country of birth of the used bulls (n = 504)

Birth year	Country[1]														Total
	AUS	AUT	BEL	CAN	CHE	CZE	DEU	DNK	FRA	GBR	HUN	IRL	ITA	NLD	USA	
1995	0	0	0	2	0	0	1	0	1	0	0	0	1	2	3	10
1996	0	0	0	1	0	0	1	1	0	1	0	0	0	0	2	6
1997	0	0	0	5	0	0	1	0	1	0	0	0	0	2	1	10
1998	1	0	0	2	0	0	2	0	0	0	0	0	1	2	4	12
1999	0	0	0	0	0	0	1	0	1	1	0	0	0	3	6	12
2000	0	1	0	2	0	0	5	0	8	1	0	0	0	5	6	28
2001	0	0	0	4	0	0	1	0	3	0	0	0	1	6	12	27
2002	0	0	0	3	1	1	3	0	0	0	2	0	4	7	4	25
2003	0	0	0	4	0	0	5	0	0	0	0	0	6	9	8	32
2004	0	0	1	2	0	0	0	0	2	1	0	0	3	10	26	45
2005	0	0	0	6	0	0	3	1	2	2	0	1	2	18	18	53
2006	0	0	1	6	0	1	5	1	2	1	0	0	10	14	21	62
2007	0	0	0	7	0	0	1	0	3	4	0	0	3	8	12	38
2008	0	0	0	3	0	0	1	0	6	2	0	0	1	20	9	42
2009	0	0	1	0	0	0	0	0	0	0	0	0	2	8	14	25
2010	0	0	3	3	0	0	2	0	0	0	0	0	1	20	3	32
2011	0	0	6	2	0	0	2	0	0	0	0	0	0	17	1	28
2012	0	0	1	0	0	0	0	0	0	0	0	0	0	15	0	16
2013	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Total	1	1	13	52	1	2	34	3	29	13	2	1	35	167	150	504

[1]: AUS: Australia; AUT: Austria; BEL: Belgium; CAN: Canada; CHE: Switzerland; CZE: Czech Republic; DEU: Germany; DNK: Denmark; FRA: France; GBR: United Kingdom; HUN: Hungary; IRL: Ireland; ITA: Italy; NLD: Netherlands; USA: United States of America.

3.3. Genetic and Phenotype Correlations

As explained previously, even if NUE and NL are somewhat related, they do not represent exactly the same biological background. Therefore, genetic and phenotypic correlations between PNUE and PNL were computed, and the results are reported in Table 3-3. As expected by their definition, PNUE and PNL showed negative correlations between them. Inside the same trait group, the correlations were positive. The genetic and phenotypic correlation ranged between −0.89 and 0.72, and −0.45 and 0.16, respectively. Additional features appeared. First correlations were, in absolute values, stronger among different traits in the same parity groups. We found that PNUE1 had the largest genetic negative correlation with PNL1 (−0.89) and the lowest with PNL2+ (−0.53). Meanwhile, PNL2+ had the lowest negative phenotypic correlation with PNUE1 (−0.11) and the largest with PNUE2+ (−0.45). This supported the hypothesis that first and later lactations are different traits. For both PNUE traits, even with a very high genetic correlation of 0.68, the phenotypic correlation between them was as low as 0.16. Similarly, for both PNL traits, the genetic correlation was 0.72, and the phenotypic correlation between them was as low as 0.13. Based on these results, we cannot say that PNUE and PNL are genetically similar traits. Each one is also slightly genetically different, comparing first and later lactations. Phenotypically, large differences also appeared among all studied traits.

Table 3-3. Genetic correlations (above the diagonal) and phenotype correlations (below the diagonal) among reported N use efficiency and loss traits in primiparous (n = 16,456) and multiparous (n = 29,707) Holstein cows

Trait[1]	PNUE1	PNL1	PNUE2+	PNL2+
PNUE1		-0.89 ± 0.04	0.68 ± 0.08	-0.53 ± 0.08
PNL1	-0.42 ± 0.01		-0.63 ± 0.08	0.72 ± 0.07
PNUE2+	0.16 ± 0.01	-0.13 ± 0.01		-0.79 ± 0.04
PNL2+	-0.11 ± 0.01	0.13 ± 0.01	-0.45 ± 0.01	

[1] PNUE1 = predicted N use efficiency in first-parity Holstein cows; PNL1 = predicted N losses in first-parity Holstein cows; PNUE2+ = predicted N use efficiency in second- to fifth-parity Holstein cows; PNL2+ = predicted N losses in second- to fifth-parity Holstein cows.

3.4. Approximate Genetic Correlations

Results obtained so far indicate that the PNUE and PNL traits show potential for genetic improvement; however, it is also necessary to assess how these novel traits correlate to other traits already under selection in most Holstein populations. Table 3-4 describes approximate genetic correlations between PNUE, PNL, and 30 traits of interest using Blanchard's approach by EBV. Estimating genetic relationships among traits under selection may reflect genetic trends for selected traits instead of real genetic correlations. Therefore, to evaluate the suitability of this approximate approach, we calculated the approximate genetic correlations of milk yield with other traits using the same approach (Table 3-4). Results showed that the approximate genetic correlations between milk yield and other traits were similar to genetic correlations found in other studies (Campos et al., 2015; Zhao et al., 2015). These results supported our hypothesis that the use of Blanchard's approach in our context can yield reasonable approximate correlations.

As shown in Table 3-4, the approximate genetic correlations of PNUE and yield traits (milk, fat, and protein) ranged from 0.04 to 0.47. In particular, milk and protein yields showed relevant positive correlations, indicating that their improvement may have increased NUE in the past, as expressed by PNUE. Milk fat percentage showed negative correlations (−0.19 and −0.14), and protein percentage nearly zero correlations, indicating that the correlations between yields and PNUE might essentially be due to the high correlations with milk yield. Results for PNL have to be considered always bearing in mind that lower PNL values are associated with better efficiency. When considering PNL, current selection favoring higher fat and protein yields, as in the Walloon selection index V€G, would increase PNL and therefore decrease N efficiency. Based on these results we can again argue that PNUE and PNL seem to be genetically very different traits.

Table 3-4. Approximate genetic correlation[1] between predicted N use efficiency and N loss, and other traits evaluated by EBV of selected bulls[2] (n = 504)

Trait	Milk yield	PNUE1	PNL1	PNUE2+	PNL2+
Yield traits					
Milk Yield	NA[3]	0.47 ± 0.06	-0.10 ± 0.07	0.21 ± 0.06	0.10 ± 0.07
Fat Yield	0.43 ± 0.04	0.24 ± 0.06	0.22 ± 0.07	0.04 ± 0.07	0.38 ± 0.06
Protein Yield	0.78 ± 0.03	0.42 ± 0.06	0.15 ± 0.06	0.24 ± 0.06	0.30 ± 0.06
Milk composition					
Fat Percent	-0.58 ± 0.04	-0.19 ± 0.08	0.28 ± 0.07	-0.14 ± 0.07	0.25 ± 0.07
Protein Percent	-0.43 ± 0.04	-0.06 ± 0.08	0.32 ± 0.07	0.03 ± 0.07	0.25 ± 0.07
Health, longevity and other functional traits					
Udder Health	-0.18 ± 0.05	-0.29 ± 0.07	0.21 ± 0.06	-0.22 ± 0.06	0.18 ± 0.06
Longevity	-0.05 ± 0.05	-0.33 ± 0.08	0.39 ± 0.06	-0.31 ± 0.06	0.36 ± 0.06
Female Fertility	-0.46 ± 0.04	-0.44 ± 0.07	0.41 ± 0.07	-0.42 ± 0.06	0.36 ± 0.06
BCS	-0.05 ± 0.05	-0.14 ± 0.07	0.41 ± 0.07	-0.12 ± 0.07	0.33 ± 0.07
Direct Calving Ease	0.20 ± 0.05	-0.07 ± 0.07	0.39 ± 0.07	-0.35 ± 0.07	0.57 ± 0.06
Maternal Calving Ease	0.03 ± 0.06	-0.34 ± 0.07	0.32 ± 0.06	-0.36 ± 0.06	0.31 ± 0.06
Conformation traits					
Stature	-0.09 ± 0.05	-0.14 ± 0.08	-0.05 ± 0.08	-0.14 ± 0.07	0.00 ± 0.07
Chest Width	0.14 ± 0.05	0.01 ± 0.07	0.16 ± 0.07	-0.04 ± 0.07	0.18 ± 0.07
Body Depth	0.05 ± 0.05	0.03 ± 0.07	-0.09 ± 0.07	-0.03 ± 0.07	-0.01 ± 0.07
Rump Angle	0.06 ± 0.05	0.14 ± 0.06	-0.16 ± 0.06	0.17 ± 0.06	-0.16 ± 0.06
Rump Width	0.06 ± 0.05	0.07 ± 0.07	-0.13 ± 0.07	0.05 ± 0.07	-0.12 ± 0.07
Foot Angle	0.08 ± 0.05	-0.09 ± 0.07	0.21 ± 0.07	-0.13 ± 0.07	0.23 ± 0.07
Rear Leg Set	-0.05 ± 0.04	-0.10 ± 0.07	-0.03 ± 0.06	-0.13 ± 0.06	0.00 ± 0.06
Rear Leg Rear View	0.02 ± 0.05	0.06 ± 0.07	-0.04 ± 0.07	0.15 ± 0.06	-0.11 ± 0.06
Udder Depth	-0.37 ± 0.04	-0.46 ± 0.06	0.18 ± 0.07	-0.36 ± 0.06	0.12 ± 0.06
Udder Support	-0.06 ± 0.05	-0.08 ± 0.07	-0.20 ± 0.07	-0.03 ± 0.07	-0.18 ± 0.07
Fore Udder	-0.29 ± 0.04	-0.29 ± 0.07	-0.06 ± 0.07	-0.11 ± 0.07	-0.16 ± 0.07
Front Teat Placement	0.07 ± 0.04	-0.11 ± 0.07	0.17 ± 0.06	-0.12 ± 0.06	0.12 ± 0.06
Teat Length	0.11 ± 0.04	0.03 ± 0.06	-0.15 ± 0.06	-0.03 ± 0.06	-0.06 ± 0.06
Rear Udder Height	0.05 ± 0.05	-0.04 ± 0.07	-0.21 ± 0.07	-0.02 ± 0.07	-0.14 ± 0.07
Rear Teat Placement	0.02 ± 0.05	0.06 ± 0.07	-0.22 ± 0.06	0.15 ± 0.06	-0.26 ± 0.06
Angularity	-0.06 ± 0.04	-0.02 ± 0.07	-0.32 ± 0.07	0.02 ± 0.07	-0.26 ± 0.07
Overall Feet and Leg Score	0.00 ± 0.05	-0.02 ± 0.07	0.01 ± 0.06	0.09 ± 0.07	-0.07 ± 0.06
Overall Udder Score	-0.23 ± 0.05	-0.34 ± 0.06	-0.01 ± 0.07	-0.23 ± 0.06	-0.04 ± 0.07
Overall Conformation Score	-0.12 ± 0.05	-0.23 ± 0.07	-0.07 ± 0.07	-0.16 ± 0.07	-0.06 ± 0.07

[1]: approximate genetic correlations were estimated using the strategy presented by Blanchard et al. (1983).

[2] PNUE1: Predicted N use efficiency in first parity Holstein cows.

[3] PNL1: Predicted N losses in first parity Holstein cows.

[4] PNUE2+: Predicted N use efficiency in second to fifth parity Holstein cows.

[5] PNL2+: Predicted N losses in second to fifth parity Holstein cows.

As explained previously, efficiency traits may be in opposition to health and other functional traits. As shown in Table 3-4, more favorable PNUE genetically decreased other important traits such as udder health, female fertility, longevity, and calving ease (direct and maternal). The effect on BCS was less clear. For most traits, the effects were as expected because of their links to milk yield, except for the effects on longevity and calving ease, which were clearly stronger than expected, indicating a specific direct link. To clarify why such strong correlations were observed for calving ease, approximate genetic correlations were computed between calving ease traits and all other traits (detailed results not shown). A common rationale is that sires transmitting bigger size also generate more direct calving difficulties. This was only partially confirmed, as correlations between direct calving ease and body development traits were only between −0.15 and −0.31. For maternal calving ease, except stature (0.22), correlations were very close to zero. However, for direct calving ease correlations with health and other functional traits were all positive and rather strong between 0.28 for udder health and 0.53 for longevity. For maternal calving ease, the corresponding correlations were all positive but lower. Except for BCS, with a correlation that was very close to zero, the correlations ranged from 0.13 for udder health to 0.27 for longevity. This finding might indicate that animals that are more efficient are unfortunately less robust and show lower longevity. Findings also indicate that all health and other functional traits are similarly affected, including calving ease traits. The results for PNL were as expected, nearly equivalent to those for PNUE except for BCS, where we observed a much stronger negative association with N efficiency, expressed through PNL. This could indicate that selecting for direct reduced N loss does not only affect N excreted to the environment but also leads to less N available to maintain body condition, probably through reduced N intake. All our results rather strongly support the hypothesis that feed efficiency traits, here N, are in opposition to health and other functional traits.

For conformation traits, correlations were only reasonably high between udder traits, or udder-related traits and PNUE. By comparing with the correlations between these traits and milk yield, it appeared that, with a high likelihood, the observed correlations are due to the strong links between milk yield and PNUE. It is difficult to interpret other correlations, most being very small. Future studies could confirm some indications, such as the negative effect on stature, which may be part of the reduction in the availability of N for body growth (Figure 3-1). In addition, cattle with large stature consume a lot of energy for maintenance. For PNL the results were rather diverse, with a less strong association with udder traits. As expected by the results from BCS, animals that were genetically more efficient for N, expressed through lower PNL values, tended to be more angular.

In short, increasing PNUE seems to have a favorable selection response for yield traits, but not favorable ones for udder health, longevity, and reproduction. According to the definition of PNL, lower PNL may be beneficial and reduce environmental pollution. By correcting for the desired direction of selection (i.e., lower PNL values), most traits (19 traits) showed the same trend in their genetic relationship with PNL as with PNUE (Table 3-4). However, even if most correlations made perfect sense, the biological links are not always clear. Moreover, we should not forget that we studied here the proxies (PNUE and PNL) and not the underlying direct traits NUE and NL. The same situation (i.e., working on proxies based on imperfect predictions and not the direct traits) has appeared in many other studies (e.g., Zaalberg et al., 2020) that estimated approximate genetic correlation between the MIR predicted trait and other traits.

However, a feature of MIR-based predictions is that, when prediction equations with improved prediction accuracy become available, we can easily recalculate improved proxy phenotypes PNUE and PNL. Because we cannot expect direct NUE or NL to become available on a large scale, we have to rely on the use of PNUE or PNL to select cows. This means that NUE or NL could become part of the breeding goal that contains the traits we want to improve. The traits PNUE or PNL would then be part of the index traits. This situation is very similar to the current selection against mastitis: the breeding goal trait is the incidence of mastitis, but the index trait is very often SCS. Furthermore, this study showed several other directions where future research is required. First, the well-known issue of using DMI versus residual feed intake parallels our situation, where the components of NUE and NL are the N content of DMI and the N content of protein yield, the difference between those (i.e., NL) mirroring residual feed intake. All of these warrants additional research. Second, the design of this study did not allow the investigation of the relationship between PNUE (PNL) and other direct measures of N (e.g., milk urea nitrogen) that might be easier to comprehend.

4. Conclusions

In summary, our study showed first estimates of genetic parameters for PNUE and PNL traits and approximate genetic correlations between PNUE, PNL, and other traits currently reported by Interbull. The estimated heritabilities of PNUE and PNL ranged from 0.12 to 0.14, and the repeatabilities ranged from 0.40 to 0.55, respectively. These results seem to suggest that selection of N efficiency-related traits is possible. Our results also showed that N efficiency traits could impair health and other functional traits when only NUE is selected. It should be noted that the data we studied were only from the early lactation (the first 50 DIM). Finally, if validated for their link to

N excretion, PNUE and PNL could also be used to aid N management in dairy farms to improve farmer's profits as well as to reduce environmental pollution.

5. Acknowledgments

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Yansen Chen. The GplusE project has received funding from the European Union's Seventh Framework Programme (Brussels, Belgium) for research, technological development, and demonstration, under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. The authors thank the 5 organizations and universities Royal Veterinary College (London, UK), Walloon Breeders Association (Ciney, Belgium), Irish Cattle Breeding Federation Society Limited (Bandon, Ireland), Huazhong Agricultural University (Wuhan, China), and CRV (Arnhem, the Netherlands) for collecting and providing the MIR data used here. The authors acknowledge the support of the Walloon Government (Service Public de Wallonie–Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE; Namur, Belgium) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech (Liège, Belgium) provided by the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the Fonds de la Recherche Scientifique–FNRS under grants no. T.0095.19 (PDR DEEP-SELECT) and J.0174.18 (CDR PREDICT-2). The authors have not stated any conflicts of interest.

6. References

Adenuga, A.H., J. Davis, G. Hutchinson, T. Donnellan, and M. Patton. 2019. Environmental efficiency and pollution costs of nitrogen surplus in dairy farms: a parametric hyperbolic technology distance function approach. Environ. Resour. Econ. 74:1273–1298. https://doi.org/10.1007/s10640-019-00367-2.

Ariyarathne, H., M. Correa-Luna, H.T. Blair, D.J. Garrick, and N. Lopez-Villalobos. 2019. Estimation of heritabilities for milk urea and efficiency of crude protein utilization by day of lactation in a New Zealand dairy cow herd. N. Z. J. Anim. Sci. Prod. 79:183–187.

Blanchard, P.J., R.W. Everett, and S.R. Searle. 1983. Estimation of genetic trends and correlations for Jersey cattle. J. Dairy Sci. 66:1947–1954. https://doi.org/10.3168/jds.S0022-0302(83)82033-5.

Campos, R.V., J.A. Cobuci, E.L. Kern, C.N. Costa, and C.M. McManus. 2015. Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. Asian-Australas J. Anim. Sci. 28:476–484. https://doi.org/10.5713/ajas.14.0288.

Cantalapiedra-Hijar, G., H. Fouillet, J.F. Huneau, A. Fanchone, M. Doreau, P. Nozière, and I. Ortigues-Marty. 2016. Relationship between efficiency of nitrogen utilization and isotopic nitrogen fractionation in dairy cows: contribution of digestion v. metabolism? Animal 10:221-229. https://doi.org/10.1017/S1751731115002025.

Cantalapiedra-Hijar, G., R.J. Dewhurst, L. Cheng, A.R.J. Cabrita, A.J.M. Fonseca, P . Nozière, D. Makowski, H. Fouillet, and I. Ortigues-Marty. 2018. Nitrogen isotopic fractionation as a biomarker for nitrogen use efficiency in ruminants: a meta-analysis. Animal 12:1827–1837. https://doi.org/10.1017/S17517311170 03391.

Canty A., and B. Ripley. 2019. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-24.

Castillo, A.R., E. Kebreab, D.E. Beever, J.H. Barbi, J.D. Sutton, H.C. Kirby, and J. France. 2001. The effect of protein supplementation on nitrogen utilization in lactating dairy cows fed grass silage diets. J. Anim. Sci. 79:247–253. https://doi.org/10.2527/2001.791247x.

Dowle, M. and A. Srinivasan (2019). data.table: Extension of `data.frame`. R package version 1.12.8. https://CRAN.R-project.org/package=data.table.

Gengler, N., M. Hostens, and GplusE Consortium. 2018. How fast can we change resilience and efficiency through breeding and management? Accessed Jan. 21, 2021. https://orbi.uliege .be/bitstream/2268/234176/2/S21_03_Gengler.pdf.

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, and 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.

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.

Hayes, J.F., and A.G.A. Ageeb. 2002. Genetic parameters for feed efficiency in Holsteins. Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. Accessed Jan. 21, 2021. http://www.wcgalp.org/system/files/proceedings/2002/genetic-parameters-feed-efficiency-holsteins.pdf.

Lopez-Villalobos, N., M. Correa-Luna, J.L. Burke, N.W. Sneddon, M.M. Schutz, D.J. Donaghy, and P.D. Kemp. 2018. Genetic parameters for milk urea concentration and milk traits in New Zealand grazing dairy cattle. N. Z. J. Anim. Sci. Prod. 78:56-61.

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.

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. http://nce.ads.uga.edu/wiki/doku.php?id=documentation.

Montaldo, H.H., H. Castillo-Juárez, M. Valencia-Posadas, E.G. Cienfuegos-Rivas, and F.J. Ruiz-López. 2010. Genetic and environmental parameters for milk production, udder health, and fertility traits in Mexican Holstein cows. J. Dairy Sci. 93:2168–2175. https://doi.org/10.3168/jds.2009-2050.

Nasrollahi, S.M., P. Nozière, R.J. Dewhurst, C. Chantelauze, L. Cheng, E. Froidmont, C. Martin, and G. Cantalapiedra-Hijar. 2019. Natural 15N abundances in plasma and urea-N concentration in milk as biomarkers of urinary N excretion in dairy cows: a meta-analysis. EAAP Scientific Series. 138:267–268. https://doi.org/10.3920/978-90-8686-891-9.

Phuong, H.N., N.C. Friggens, I.J.M. de Boer, and P. Schmidely. 2013. Factors affecting energy and nitrogen efficiency of dairy cows: a meta-analysis. J. Dairy Sci. 96:7245–7259. https://doi.org/10.3168/jds.2013-6977.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Vanlierde, A., M.-L. Vanrobays, F. Dehareng, E. Froidmont, H. Soyeurt, S. McParland, E. Lewis, M.H. Deighton, F. Grandl, M. Kreuzer, B. Gredler, P. Dardenne, and N. Gengler. 2015. Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. J.

Dairy Sci. 98:5740–5747. https://doi.org/10.3168/jds.2014-8436.

Wheadon, N.M., M. McGee, G.R. Edwards, and R.J. Dewhurst. 2014. Plasma nitrogen isotopic fractionation and feed efficiency in growing beef heifers. Br. J. Nutr. 111:1705–1711. https://doi.org/10.1017/S0007114513004078.

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.

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Zaalberg, R. M., A. J. Buitenhuis, U. K. Sundekilde, N. A. Poulsen, and H. Bovenhuis. 2020. Genetic analysis of orotic acid predict-ed with Fourier transform infrared milk spectra. J. Dairy Sci. 103:3334–3348. https://doi.org/10.3168/jds.2018-16057.

Zhao, F.P., G. Guo, Y.C. Wang, X.Y. Guo, Y. Zhang, and L.X. Du. 2015. Genetic parameters for somatic cell score and production traits in the first three lactations of Chinese Holstein cows. J. Integr. Agric. 14:125–130. https://doi.org/10.1016/S2095-3119(14)60758-9.

4

Chapter IV Defining a nitrogen efficiency index in Holstein cows and assessing its potential impact on the breeding program of bulls

Adapted from: Chen, Y., H. Atashi, C. Grelet, S. Vanderick, H., Hu, and N. Gengler. 2022. Defining a nitrogen efficiency index in Holstein cows and assessing its potential impact on the breeding program of bulls. J. Dairy Sci. 105: 7575-7587. https://doi.org/10.3168/jds.2021-21681.

Foreword

Genetic analyses of the traditional and novel proxies of NUE are shown in Chapters 2 and 3, respectively. However, the genetic correlations between milk urea concentration (MU), predict nitrogen use efficiency (PNUE) and nitrogen losses (PNL) are very low (shown in the 1.3 section of Chapter 7). In this Chapter, three NUE-related features were integrated into the nitrogen efficiency index (NEI), which considered both NUE and N pollution at the same time. Furthermore, the impacts of adding NEI to the genetic selection program in the Walloon Region of Belgium were explored.

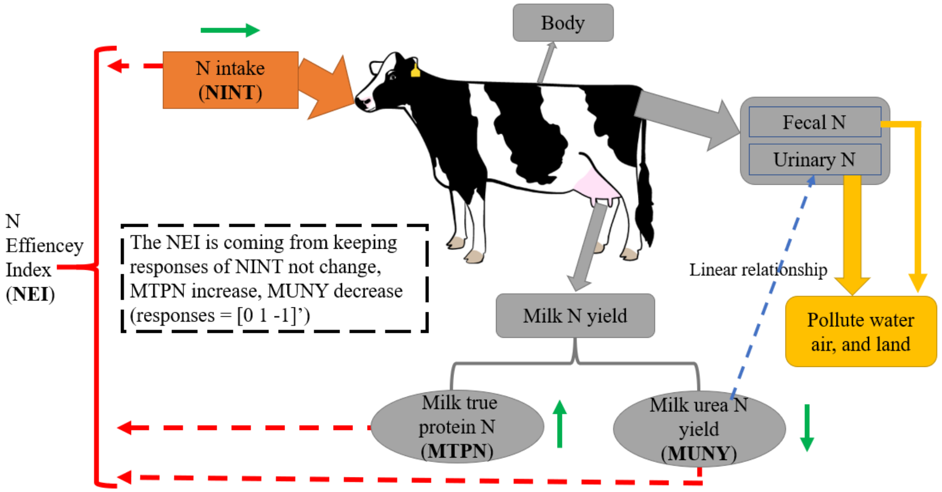
Abstract

The purposes of this study were to 1) explore the relationship between three milk mid-infrared (MIR) predicted features including nitrogen (N) intake (NINT), milk true protein N (MTPN), and milk urea N yield (MUNY); 2) integrate these three features into an N efficiency index (NEI) and analyses approximate genetic correlations between the NEI and 37 traits (indices) of interest; 3) assess the potential impact of including the NEI into breeding programs of bulls. The edited data were 1,043,171 test-day records on 342,847 cows in 1,931 herds and 143,595 test-day records on 53,660 cows in 766 herds used for estimating breeding values (EBV) and variance components, respectively. The used records were within 5 to 50 days in milk (DIM). The records were grouped into primiparous and multiparous. The genetic parameters for the included MIR features and EBV of the animals included in the pedigree were estimated using a multiple-trait repeatability animal model. Then, the EBV of the NINT, MTPN, MUNY were integrated into the NEI using a selection index assuming weights based on the N partitioning. The approximate genetic correlations between the NEI and 37 traits (indices) of interest were estimated using the EBV of the selected bulls. The bulls born from 2011 to 2014 with NEI were selected and the NEI distribution of these bulls having EBV for the eight selected traits (indices) was checked. The heritability and repeatability estimated for NINT, MTPN, and MUNY ranged from 0.09 to 0.13, and 0.37 to 0.65, respectively. The genetic and phenotypic correlations between NINT, MTPN, and MUNY ranged from −0.31 to 0.87, and −0.02 to 0.42, respectively. The NEI ranged from −13.13 to 12.55 kg/d. In total, 736 bulls with reliability ≥ 0.50 for all included traits (NEI and 37 traits or indices) and at least 10 daughters distributed in at least 10 herds were selected to investigate genetic aspects of the NEI. The NEI had positive genetic correlations with production yield traits (0.08 to 0.46), and negative genetic correlations with the investigated functional traits and indices (−0.71 to −0.07), except for the production economic index and functional type economic index. The daughters of bulls with higher NEI had lower NINT and MUNY, and higher MTPN. Furthermore, 26% of the bulls (n = 50) with NEI born between 2011 to 2014 had higher NEI and global economic index than the average in the selected bulls. Finally, the developed NEI has the advantage of large-scale prediction and therefore has the potential for routine application in dairy cattle breeding in the future.

Key words: N intake, genetic correlation, health, mid-infrared spectrum

1. Introduction

The economic importance of genetic improvement for efficiency traits in cattle is recognized by the world's cattle producers (Brito et al., 2020). Measuring and improving efficiency is not only beneficial to the protection of the environment, but also beneficial to the farm, can promote the sustainable economic development of dairy production (Chen et al., 2021c), and strengthen its social acceptability. In the process of studying efficiency, researchers put forward various indicators representing dairy cow efficiency which can be defined very broadly. In the context of direct efficiency of a given animal, efficiency is often linked to feeding efficiency (FE) that can be divided further, i.e. into energy and nitrogen efficiency. Often FE is associated with the energy (e.g., energy intake (EI), energy balance (EB)) and an expression as residual trait compared to expectations, e.g. residual feed intake (RFI), is used. Many works for FE and energy have been done (McParland et al., 2015; Brito et al., 2020), especially FE has started being used in the dairy cattle breeding system in some countries (e.g., Australia, the US) (Pryce et al., 2014; Parker Gaddis et al., 2021). However, there are relatively few studies on N use efficiency (NUE) of dairy cows, except for those addressing milk urea concentration (MU) or milk urea nitrogen (MUN) as implicit proxies (Bobbo et al., 2020; Chen et al., 2021b). To our knowledge, only we have recently performed the genetic analyses of predicted NUE for dairy cows in early lactation (Chen et al., 2021c). Moreover, the NUE by cattle is very low (20% to 44%) and could be improved (Cantalapiedra-Hijar et al., 2018; Grelet et al., 2020). For example, recently Spanghero and Kowalski (2021) conducted a meta-analysis of 86 N balance experiments carried out in the past 20 years, and found that the average NUE in cattle was 27% and more than half of the feed N is excreted through urine and feces.



The NEI is coming from keeping responses of NINT not change, MTPN increase, MUNY decrease (responses = [0 1 -1]')

Figure 4-1. A new N efficiency index was built based on EBV from N intake, milk true protein N, and milk urea N yield.

As often for novel traits, the definition of NUE can be very different. Based on Calsamiglia et al. (2010), Grelet et al. (2020) defined NUE as milk N divided by the N intake (NINT). This definition has several shortcomings. First, from a mathematical point of view changes in ratios are highly unpredictable as they can come from changes in the denominator or the numerator. Moreover, as mentioned by Grelet et al. (2020), this definition for NUE does not account for the actual N losses through urine and feces, making it impossible to calculate the N balance. Consequently, early lactation cows having limited intake capacities and producing high quantities of milk may experience a negative N balance on top of the negative energy balance. Looking only at N intake and milk N, the NUE would be artificially high. There is then a risk of confusing artificially high NUE with negative N balance, and trying to improve NUE may increase the difficulty induced by severe mobilization of body reserves. Additionally, our previous study found that the genetic correlation between predicted NUE and MU was close to 0 (Chen et al., 2021a). Based on the above reasons, we aimed to build a new N efficiency index (NEI) considering the NUE and N losses at the same time to avoid these shortcomings (Figure 4-1). The NEI and its composition both differ from the predicted NUE trait of our previous study. The MUN yield (MUNY) is linearly proportional to the urinary urea-N excretion when defined as a quantity excreted (Wisconsin Alumni Research Foundation, 2015). So the NEI index can be built based on NINT, milk true protein N (MTPN), and MUNY. Among these three features, the MTPN and MUNY can be easily measured using traits that are currently recorded as explained later. However, the NINT is a feature that is difficult and expensive to measure in routine. The composition of milk is affected by the animal diet, and milk mid-infrared (MIR) spectra can reflect the changes in milk composition, which suggests that MIR can indirectly reflect the composition of the diet. (Klaffenböck et al., 2017). Grelet et al. (2020) developed a NINT prediction equation for dairy cows based on milk MIR spectra which have been already applied to the Walloon Region of Belgium dataset (Chen et al., 2021a). The MIR spectra is an inexpensive method for predicting features, and it has been applied to predict various traits in dairy cows (Grelet et al., 2021). At the same time, if a new trait (or index) is added to the breeding program, it is necessary to know its potential effects on other traits included in the selection index and proxies of NUE (e.g., MU).

The purposes of this research were to 1) estimate the genetic parameters and EBV of NINT, MTPN, and MUNY; 2) integrate these three features into an NEI based on the selection index and investigate the approximate genetic correlations between the NEI and 37 traits (indices) of interest based on EBV of selected bulls; and 3) assess the potential impact of including the NEI into breeding programs of bulls.

2. Materials and methods

2.1. Data

All data were collected from 2001 to 2019 as the official milk recording in the Walloon Region of Belgium. The milk samples were analyzed by mid-infrared spectrometry (commercial instruments from FOSS) to assess milk compositions (among which MU) and generate MIR spectra. Daily milk yield (MY), protein percentage (PP), and MU were restricted to the range of 1 to 90 kg, 1 to 7%, and 2 to 70 mg/dL, respectively. The filtering procedure used for NINT was the same as that explained by Chen et al. (2021c). In brief, two criteria are considered in the filtering procedure: 1) the standardized Mahalanobis distance between the new predicted spectrum and the calibration dataset is ≤ 3; 2) the predicted value of NINT is within the range of 3 standard deviations (SD) of the mean. The formulas used to compute MTPN = [(MY × PP / 6.38) – MUNY]; and MUNY = [(MU / 2.14) × MY] (WHO and FAO, 2011) allowed to transform the concentration to yield. Furthermore, the range of days in milk (DIM) was restricted between 5 and 50, lactation number was restricted between first and five, and calving age was restricted between 22 to 39, 34 to 53, 47 to 68, 59 to 83, and 71 to 98 months for the first to fifth parity, respectively. At least two of the three included features had to be available. Finally, 1,043,171 test-day records on 342,847 cows in 1,931 herds collected between 2001 and 2019 remained. The pedigree related to the dataset comprised 504,333 animals (17,573 males). Besides, each feature was divided into primiparous (n = 404,312) and multiparous (n = 638,859) (including second to fifth parity) classes. Hereafter, the features for NINT, MTPN, and MUNY are identified as NINT1, MTPN1, MUNY1 for primiparous cows, and NINT2+, MTPN2+, and MUNY2+ for multiparous cows.

Milk MIR spectra of 53,660 cows in 776 herds were used leading to 143,595 NINT records. All milk MIR spectra were standardized according to Grelet et al. (2015). The NINTs of each cow were predicted by the equation developed based on the models established by Grelet et al. (2020) using milk MIR spectra, MY, and parity as additional predictors, through support vector machine regression. The determination coefficient (R^2) and root mean square error (RMSE) of validation of the NINT model were 0.71 and 0.07 kg/d, respectively.

2.2. (Co)variance Components and Genetic Parameters

In total, 143,595 records (NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+) on 53,660 cows were used to estimate (co)variance components. The used pedigree included 132,056 animals (7,340 males). A six-trait (six traits including three features in two parity classes) repeatability animal model was used to estimate the (co)variance components. The used model was based on that presented by Chen et al. (2021c) to a different set of six traits:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Xb} + \mathbf{Qq} + \mathbf{W1c} + \mathbf{W2p} + \mathbf{Za} + \mathbf{e} \qquad [1]$$

where **y** was a vector of NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+. In each trait, all effects in this model were the same as Chen et al. (2021c). In brief, **h** was a vector of fixed effect of Herd-Year-Season of calving; **b** was a vector of fixed regression coefficients for standardization DIM and its quadratic; **q** was a vector of fixed regression coefficients of the standardization age of calving, defined as a constant, linear and quadratic regression defined inside parities (from first to fifth parity); **c** was a vector of non-genetic cow random effect; **p** was a vector of non-genetic cow × parity random effect, this effect was only modeled for NINT2+, MTPN2+, and MUNY2+, as they allowed to distinguish records for the same cow but occurring during different parities (from second to fifth parity); **a** was a vector of the random additive genetic effect; **e** was a vector of random residual effect. In addition, **H**, **X**, **Q**, **W1**, **W2**, and **Z** were incidence matrices assigning observations to effects.

The expected values and variances in Equation [1] were defined similarly to Chen et al. (2021c) but applied to a different set of six traits. The matrices of V(**c**) and V(**a**) both contained a block of 6 × 6 (co)variance matrices. For V(**e**), the diagonal and off-diagonal elements of the matrix were non-zero and zero, respectively. For V(**p**) the elementary (co)variance matrix was reduced to a 3 × 3 matrix, because only the (co)variances associated with NINT2+, MTPN2+, and MUNY2+ were present.

All computations were performed in the BLUPF90 programs (Misztal et al., 2018). The (co)variance components for NINT, MTPN, and MUNY were estimated by Gibbs sampling in THRGIBBS1F90 (version 2.118) through Equation [1], and posterior convergence was analyzed by POSTGIBBSF90 (version 3.14). Among them, the posterior means of (co)variances, heritabilities (h^2), repeatability (REP), genetic and phenotypic correlations were obtained using 40,000 samples, which is from a single chain of 220,000 after a burn-in of 20,000, and 1 in every 5 samples was saved. The formulas used to calculate the later four parameters were the same as those used by Chen et al. (2021c). Approximate standard errors (SE) of all calculated parameters were obtained using the POSTGIBBSF90 (version 3.14) program.

2.3. Estimated Breeding Value

A total of 1,043,171 records were used to estimate EBV for the included six traits based on the estimated (co)variance components. The EBVs of the included traits were estimated according to a precondition conjugate algorithm implemented in the BLUP90IOD2 (version 3.81, http://nce.ads.uga.edu/wiki/doku.php?id=readme.pcg2) program using Equation [1]. Multi-trait reliability (REL) of EBV could not be calculated directly for the six included traits, therefore we divided the six included traits into two-trait models (all effects were the same as Equation [1]) according to NINT, MTPN, and MUNY. This allows calculating the REL of EBV separately by direct inversion of the coefficient matrix for the NINT, MTPN, and MUNY. The two-trait models were then solved using the BLUPF90 (version 1.70) program to extract diagonal elements of the inverted coefficient matrix allowing to compute REL in this way hereafter called $\mathrm{REL_S}$. From this point, selection index theory was used based on a strategy proposed by VanRaden et al. (2018) to calculate REL. First, we restore the six-trait REL as $\mathrm{REL_M}$, the $\mathrm{REL_M}$ assuming that the six-trait solutions could have been obtained as index traits combing the two-trait solutions. Needed regression coefficients were calculated using selection index theory:

$$\mathbf{B} = \mathbf{G}\mathbf{G}^{*-\mathbf{1}}$$

where $\mathbf{G}$ was the full 6 × 6 matrix of the estimated genetic (co)variance of the explained six-trait model, $\mathbf{G}^*$ was equal to the $\mathbf{G}$ excluding the covariances between NINT, MTPN, and MUNY that were put to zero, and $\mathbf{B}$ was a 6 × 6 matrix of regression coefficients. Reliabilities of the six solutions were computed as the ratios between explained variances and total or maximum variances. Maximum (co)variances of six-trait solutions were computed as follows:

$$\mathbf{V_M} = \mathbf{BGB'}$$

In the next step, for each animal i, the matrix $\mathbf{R}_i$ was obtained by multiplying each element of $\mathbf{B}$ by the square root of the relevant element of $\mathrm{REL_S}$ for this animal. This allowed us to compute the (co)variances explained for this specific animal:

$$\mathbf{V}_{\mathbf{S}i} = \mathbf{R_i G R_i}'$$

Finally, only the variance ratios being relevant the following formula was used to obtain $\mathrm{REL}_{\mathrm{M}}$ for animal *i*.

$$\mathbf{REL}_{\mathbf{M}i} = \mathbf{diag}(\mathbf{V}_{\mathbf{S}i})\mathbf{diag}(\mathbf{V}_{\mathbf{M}})^{-1}$$

2.4. Nitrogen Efficiency Index

The NEI was calculated as follows:

$$\mathbf{NEI} = \mathbf{a}'\mathbf{u}$$

where **u** was a vector of EBVs for the six included traits (NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+), and **a** was a vector of relative weights. Selection index theory for desired selection response was used to obtain the **a** coefficients by computing $\mathbf{a} = \mathbf{G_C}^{-1}\mathbf{r}_{\mathbf{org}}$, where $\mathbf{G_C}$ was the 6 × 6 matrix of the genetic correlations for the six included traits, and $\mathbf{r}_{\mathbf{org}}$ was the selection response vector of the six included traits. We assumed that the selection responses (**r**) for NINT, MTPN, and MUNY were 0, 1, –1 (**r** = [0 1 –1]'), respectively, which means keeping NINT unchanged, increasing MTPN, but decreasing MUNY. All traits were expressed in the same unit (kg/d); therefore, selection responses were defined in the original scale. As the **r** was defined for the three combined features across the 2 parity classes, a transformation matrix **T** was needed to convert $\mathbf{G_C}$ to $\mathbf{G_{Ct}}$.

$$\mathbf{T} = \begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

$$\mathbf{G_{Ct}} = \mathbf{T}\mathbf{G_C}\,\mathbf{T}'$$

$$\mathbf{a}_{\mathrm{t}} = \mathbf{r}\mathbf{G}_{\mathbf{Ct}}^{-1}$$

$$\mathbf{a} = \mathbf{a}_{\mathrm{t}}\mathbf{T}$$

The REL of NEI was calculated using the method given by VanRaden et al. (2018):

$$\mathbf{REL_{NEI}} = \frac{\mathbf{w}\mathbf{G_C}\mathbf{w}'}{\mathbf{a}\mathbf{G_C}\mathbf{a}'}$$

where $\mathbf{REL_{NEI}}$ was the REL of NEI, **w** was obtained by multiplying the elements of **a** by the square root of $\mathrm{REL}_{\mathrm{M}}$ for the EBV of the six traits.

When we define **P** as the 6 × 6 matrix of the summed (co)variances for the six traits, $\mathrm{h}^2_{\mathrm{NEI}}$ was the h^2 of NEI can be computed as:

$$\mathrm{h}^2_{\mathrm{NEI}} = \frac{\mathbf{aGa}'}{\mathbf{aPa}'}$$

where **a** and **G** were explained previously.

2.5. Approximate Genetic Correlations

The approximate genetic correlations between the NEI and 37 traits (indices) of interest were estimated using the procedure proposed by Blanchard et al. (1983) based on the EBV of selected bulls. Those bulls that had NEI and had at least 10 daughters distributed in at least 10 herds, and their reliability of EBV for the NEI and 37 traits (indices) was at least 0.50, were selected. The SE of the estimated approximate genetic correlations was estimated using 1,000 bootstrapped replicates (Chen et al., 2021c). At the same time, the selected bulls were used to perform the same computation for the six traits included in the NEI; however, the reliability of the EBV of these six traits was required to be greater than 0.30 (Chen et al., 2021c).

In this study, we used EBV and its REL of 36 traits or indices (except for the MU) obtained by the national genetic evaluation system of dairy cows in the Walloon Region of Belgium in April 2021 (published in https://www.elinfo.be/telechargerEN.html). The model and algorithm used for calculating EBV of the 36 traits (indices) of interest can be found in documents (Belgium) submitted to INTERBULL (https://interbull.org/ib/geforms) or reported by Vanderick et al. (2020). The EBV of MU and its REL were reported by Chen et al. (2021b), and we used the average daily EBV of the first three parities, which was expressed on a standardized scale with a mean of 100 and an SD of 10.

The 37 traits (indices) of interest were as follows: MU, MY, fat yield (FY), protein yield (PY), fat percentage (FP), PP, udder health (UH, which represents the opposite somatic cell score), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index, member economic index, capacity economic index, udder economic index, functional type economic index, functional economic index, global economic index, stature, chest width, body depth, rump angle, rump width, foot angle, rear leg set, rear leg rear view, udder depth, udder support, fore udder, front teat placement, teat length, rear udder height, rear teat placement, angularity, overall feet and leg score, overall udder score, and overall conformation score (OCS). The definitions of all indices and their proportions to the V€G have been explained by Vanderick et al. (2020).

2.6. Potential Impact of the Selected NEI in Bull

Two different approaches (phenotypes and genetic values) were taken to check the potential impact and possibility of selecting NUE in Bull.

In the first approach, we checked the phenotypes of cows. Based on EBV computed in the previous section, bulls with bottom 5% and top 5% NEI values were selected. Based on the newly selected bulls, their daughters were screened. The means and SD for the traits (MY, PP, MU, NINT, MTPN, MUNY) of these daughters were obtained by groups of bulls. The *t*-test was used to detect differences between the 2 groups for these 6 traits.

In the second approach, we showed the distribution between NEI and 8 traits or indices in bulls. Bulls born from 2011 to 2014 with reliability ≥ 0.50 for 8 selected traits or indices (MU, MY, PY, UH, and member, udder, functional, or global economic indices) from 37 traits (indices) and at least 10 daughters distributed in at least 10 herds were selected. The distribution between the NEI of the selected bulls and the new 8 selected traits (indices) was investigated. The paired trait distribution map was divided into 4 regions based on the average value of the traits (indices) of the selected bulls. The bulls that fall into the upper right corner are considered to be better bulls, except for MU (needed in the bottom right).

To make the NEI comparable to the selected traits (indices), NEI was standardized as follows:

$$\mathrm{NEI}_{is} = \frac{\mathrm{NEI}_{i} - \mathrm{Mean}_{2015}}{\mathrm{SD}_{2015}}$$

where NEI_{is} and NEI_{i} were, respectively, the standardized and not-standardized NEI of *i* individual, Mean_{2015} and SD_{2015} were the average and SD of NEI from the cows born in 2015 (n = 17,597), respectively. Then, the NEI_{is} was expressed on a standardized scale with a mean of 100 and an SD of 10, which is the same as the LONG and other included functional traits (Vanderick et al., 2020). Additional data preparation and processing were done using R (R Core Team, 2021).

3. Results and discussion

3.1. Descriptive Statistics

The descriptive statistics of the traits involved in this study are presented in Table 4-1. The average, SD, and coefficient of variation (CV) of the included six traits in the primiparous were lower than those in the multiparous cows (parity2+), except for the CV of NINT. The MUNY has the largest CV in primiparous and multiparous. The average of MTPN (0.13 kg/d) is about one-third of NINT (0.42 kg/d), which indicates that the NUE of the animals included in this research is about 33%. This value is within the range of 20% to 44% reported by Cantalapiedra-Hijar et al. (2018) and Grelet et al. (2020).

Table 4-1. Mean, SD, and CV of related features in genetic evaluation and (co)variance component estimation data sets

Parity class[1]	MY[2] (kg/d)	PP (%)	MU (mg/dL)	NINT (kg/d)	MTPN (kg/d)	MUNY (kg/d x 1000)
Genetic evaluation dataset (total n = 1,043,171)						
1 (n = 404,312)						
Mean	25.46	3.14	22.92	/[5]	0.12	2.73
SD[3]	5.72	0.33	8.42	/	0.03	1.17
CV[4]	22.49	10.40	36.75	/	22.68	43.05
2+ (n = 638,859)						
Mean	33.37	3.23	22.84	/	0.16	3.57
SD	8.11	0.39	8.85	/	0.04	1.63
CV	24.30	12.14	38.74	/	23.81	45.82
(Co)variance component estimation dataset (total n = 143,595)						
1 (n = 44,321)						
Mean	26.60	3.15	23.40	0.42	0.13	2.92
SD	5.43	0.31	7.21	0.06	0.03	1.08
CV	20.41	9.75	30.81	14.84	20.31	36.99
2+ (n = 99,274)						
Mean	35.50	3.20	22.30	0.49	0.17	3.73
SD	7.80	0.36	7.87	0.07	0.04	1.57
CV	21.97	11.19	35.29	14.27	21.74	42.09

[1] Parity class: the parities (from 1 to 5) were divided into 2 classes based on primiparous (class 1) and multiparous (class 2+).

[2] MY - Milk Yield; PP - Protein percentage; MU - Milk urea concentration; NINT - N intake; MTPN - Milk true protein N; MUNY – Milk urea N yield

[3] SD: standard deviation

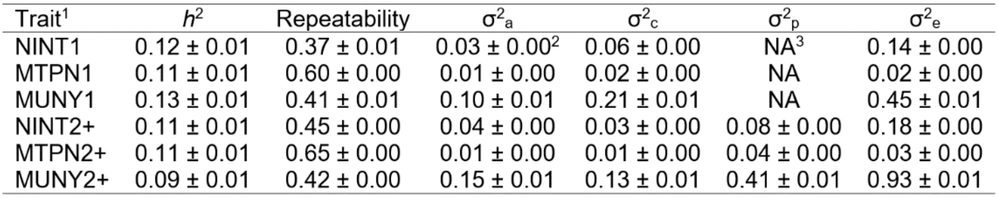
[4] CV: coefficient of variation

[5] : NINT has 143,595 records

3.2. Genetic Parameter for Six Traits

The variance components, h^2, and REP estimated for NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+ are presented in Table 4-2. The h^2 and REP of the six included traits ranged from 0.09 to 0.13 and 0.37 to 0.65, respectively. The genetic variances of the six included traits explained 12.30%, 10.52%, 12.95%, 11.21%, 11.48%, and 9.22% of the corresponding total variances. For the six traits examined, as far as we know, this is the first report on their h^2 and rep. We compared the genetic parameters of these traits with those used for the efficiency of different nutrition factors. Compared with MUNY, the MU(N) has higher h^2 and REP (Bobbo et al., 2020; Chen et al., 2021b). The NINT is similar to EI (energy efficiency) and DMI (FE). McParland et al. (2015) showed that h^2 and rep of EI predicted by MIR spectroscopy were 0.20 and 0.33, respectively. Li et al. (2016) reported that h^2 and REP of DMI in the first four weeks of lactation in Holstein cows were 0.26 and 0.68, respectively. The estimated h^2 of NINT was lower than that reported for EI and DMI, but the REP was somewhere in between.

Table 4-2. Heritability (h^2), repeatability, additive genetic variance (σ^2_a), across-parity permanent environment (non-genetic cow) variance (σ^2_c), within parity permanent environment (non-genetic cow x parity) variance (σ^2_p) (only for second and later lactations), and residual variance (σ^2_e) of the proxies for predicted N intake (NINT, expressed as 100 g/day), milk true protein nitrogen (MTPN, expressed as 100 g/day), and yield of milk urea N (MUNY, g/day) in primiparous (n = 44,321) and multiparous (n = 99,274) Holstein cows



Trait[1]	h^2	Repeatability	σ^2_a	σ^2_c	σ^2_p	σ^2_e
NINT1	0.12 ± 0.01	0.37 ± 0.01	0.03 ± 0.00[2]	0.06 ± 0.00	NA[3]	0.14 ± 0.00
MTPN1	0.11 ± 0.01	0.60 ± 0.00	0.01 ± 0.00	0.02 ± 0.00	NA	0.02 ± 0.00
MUNY1	0.13 ± 0.01	0.41 ± 0.01	0.10 ± 0.01	0.21 ± 0.01	NA	0.45 ± 0.01
NINT2+	0.11 ± 0.01	0.45 ± 0.00	0.04 ± 0.00	0.03 ± 0.00	0.08 ± 0.00	0.18 ± 0.00
MTPN2+	0.11 ± 0.01	0.65 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.04 ± 0.00	0.03 ± 0.00
MUNY2+	0.09 ± 0.01	0.42 ± 0.00	0.15 ± 0.01	0.13 ± 0.01	0.41 ± 0.01	0.93 ± 0.01

[1] Trait: NINT1 - N intake in primiparous cows; MTPN1 – Milk true protein N in primiparous cows; MUNY1 - Milk urea N yield in primiparous cows; NINT2+ - N intake in multiparous cows; MTPN2+ - Milk true protein N in multiparous cows; MUNY2+ - Milk urea N yield in multiparous cows

[2]: Standard error is less than 0.005

[3] NA: not applicable

Genetic and phenotypic correlations among the six included traits are presented in Table 4-3. As expected, the same feature has high genetic correlations between primiparous and multiparous (0.82 to 0.89). There were negative genetic correlations between NINT and MUNY (−0.31 to −0.16), and moderate positive genetic correlations were found between NINT and MTPN (0.40 to 0.51). The MTPN and MUNY showed moderate positive genetic correlations (0.37 to 0.54). However, the phenotypic correlations found between different features in the same parity class were relatively strong compared to those found for the same features between parity classes. The phenotypic correlations between NINT and MTPN in the primiparous and multiparous classes were, respectively, 0.34 and 0.42, indicating that genetic and phenotype correlations between NINT and MTPN are high. The phenotypic correlation estimated between MTPN and MUNY in the primiparous and multiparous were 0.38 and 0.41, respectively. The phenotypic correlations estimated between the remaining trait pairs were relatively low (–0.02 to 0.20).

Table 4-3. Genetic correlations (above the diagonal) and phenotypic correlations (below the diagonal) among reported N-related traits in primiparous (n = 44,321) and multiparous (n = 99,274) Holstein cows

Trait[1]	NINT1	MTPN1	MUNY1	NINT2+	MTPN2+	MUNY2+
NINT1		0.48 ± 0.04	-0.31 ± 0.06	0.89 ± 0.03	0.40 ± 0.05	-0.25 ± 0.05
MTPN1	0.34 ± 0.00[2]		0.54 ± 0.05	0.51 ± 0.05	0.82 ± 0.03	0.46 ± 0.05
MUNY1	0.06 ± 0.01	0.38 ± 0.00		-0.16 ± 0.06	0.37 ± 0.05	0.87 ± 0.02
NINT2+	0.17 ± 0.01	0.13 ± 0.01	0.01 ± 0.01		0.49 ± 0.03	-0.19 ± 0.05
MTPN2+	0.11 ± 0.01	0.20 ± 0.01	0.11 ± 0.01	0.42 ± 0.00		0.50 ± 0.04
MUNY2+	-0.02 ± 0.01	0.09 ± 0.01	0.15 ± 0.01	0.14 ± 0.00	0.41 ± 0.00	

[1] Trait: NINT1 - N intake in primiparous cows; MTPN1 – Milk true protein N in primiparous cows; MUNY1 - Milk urea nitrogen yield in primiparous cows; NINT2+ - N intake in multiparous cows; MTPN2+ - Milk true protein N in multiparous cows; MUNY2+ - Milk urea nitrogen yield in multiparous cows

[2]: Standard error is less than 0.005

3.3. Nitrogen Efficiency Index

The h^2 of NEI is low (0.06), which is similar to that estimated for the predicted NUE (Chen et al., 2021c). It has been reported that h^2 for RFI, as an indicator of FE, in the first 60 DIM was 0.10 (Jamrozik et al., 2020) and that reported for from 50 to 250 DIM was 0.14 (Li et al., 2020). It has been shown that h^2 of RFI varied in different herds and countries (Tempelman et al., 2015). The NEI and its REL ranged from – 13.13 to 12.55 kg/d, and 0.00 to 0.95. The mean REL of NEI of all animals included in the pedigree was 0.21 (SD = 0.11), which is similar to the REL of EBV of animals with RFI phenotype (Li et al., 2020). Li et al. (2020) evaluated the average REL of EBV for RFI for animals with genotypes and without phenotypes was only 0.13 (n = 1.6 million).

Although NEI has a low average REL, repeated predictions can be done cheaply. In this way, the REL of the NEI for bulls having more daughters can be increased. For example, the average REL of NEI for selected bulls in this study (n = 736) used for estimating the approximate genetic correlation was 0.68 (Figure 4-2).

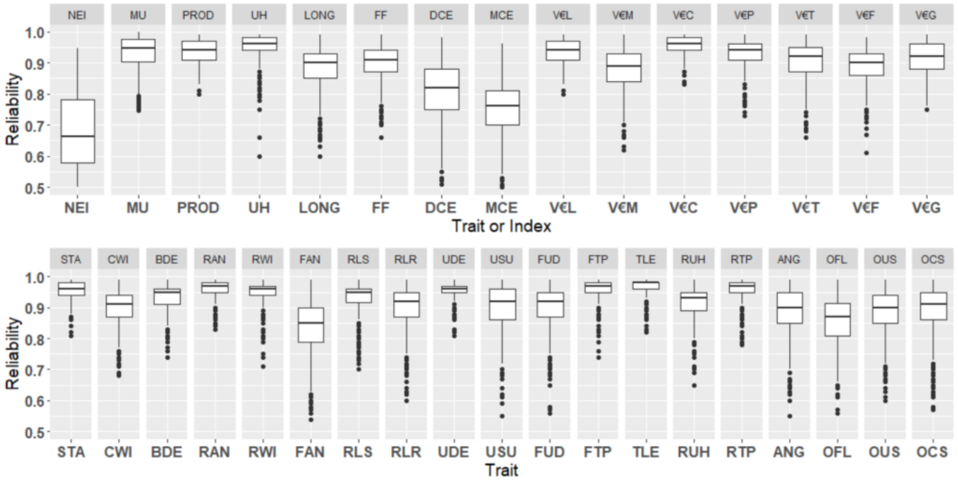


Figure 4-2. Distribution of reliability for bulls (n = 736) in nitrogen efficiency index (NEI), milk urea concentration (MU), production (PROD), udder health (UH), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index (V€L), member economic index(V€M), capacity economic index (V€C), udder economic index (V€P), functional type economic index (V€T), functional economic index (V€F), global economic index (V€G)), stature (STA), chest width (CWI), body depth (BDE), rump angle (RAN), rump width (RWI), foot angle (FAN), rear leg set (RLS), rear leg rear view (RLR), udder depth (UDE), udder support (USU), fore udder (FUD), front teat placement (FTP), teat length (TLE), rear udder height (RUH), rear teat placement (RTP), angularity (ANG), overall feet and leg score (OFL), overall udder score (OUS), and overall conformation score (OCS). (Note: production includes milk yield, fat yield, protein yield, fat percentage, protein percentage)

3.4. Approximate Genetic Correlations between the NEI and 37 Traits (indices) of interest

In total, 736 bulls with REL ≥ 0.50 for NEI and 37 traits (indices) of interest and at least 10 daughters distributed in at least 10 herds were selected for estimating the approximate genetic correlations between the NEI and the examined traits of interest. The distribution of the REL of the examined traits (indices) for the selected bulls is presented in Figure 4-2 and Table 4-1S (appendix). The average RELs of all traits (indices) in the selected bulls were greater than 0.63, which guaranteed the reliability of our subsequent results. The countries of origin of the most bulls were the US (235), Canada (130), and the Netherlands (123).

The estimated approximate genetic correlations between NEI and the included traits (indices) are shown in Figure 4-3, and will be referred to as genetic correlations in the latter part of this article. As expected, the NEI and MU had a strong negative genetic correlation. This shows that increased NEI is associated with increased NUE and decreased N pollution. Positive genetic correlations were observed between NEI and yield traits (MY, FY, and PY), ranging from 0.08 to 0.46. The genetic correlation between NEI and FP was negative, but that found between NEI and PP was positive. This shows that an increased NEI is associated with increased production of N (MY, and PY), a decreased FP, and an increased PP. This finding is consistent with the genetic relationships reported between predicted NUE and production by Chen et al (2021c). This shows that in the current breeding system in the Walloon Region of Belgium, increasing production traits will increase the NUE of cattle when other traits are not considered. The former studies also proved that cows with a higher FE had a higher milk production (Vallimont et al., 2011; Köck et al., 2018).

Unfavorable genetic correlations were observed between NEI and UH (–0.27). The predicted NUE and UH were reported to have a negative genetic correlation (Chen et al., 2021c). This means that efficiency traits and UH are negatively genetically correlated. Increased NEI led to increased milk production, subsequently increased the intensity of udder use, and reduced its health. The NEI had negative genetic correlations with LONG (–0.21) and FF (–0.24). The previous study found that cows with a higher FE had a longer day open (support our results) and a longer production life (different from our results) (Vallimont et al., 2013). Unfavorable genetic correlations were found between the NEI and calving ease traits (DCE, MCE). One explanation is that NEI and yield traits (MY, FY, and PY) were positively genetically correlated, but yield traits were negatively correlated with DCE.

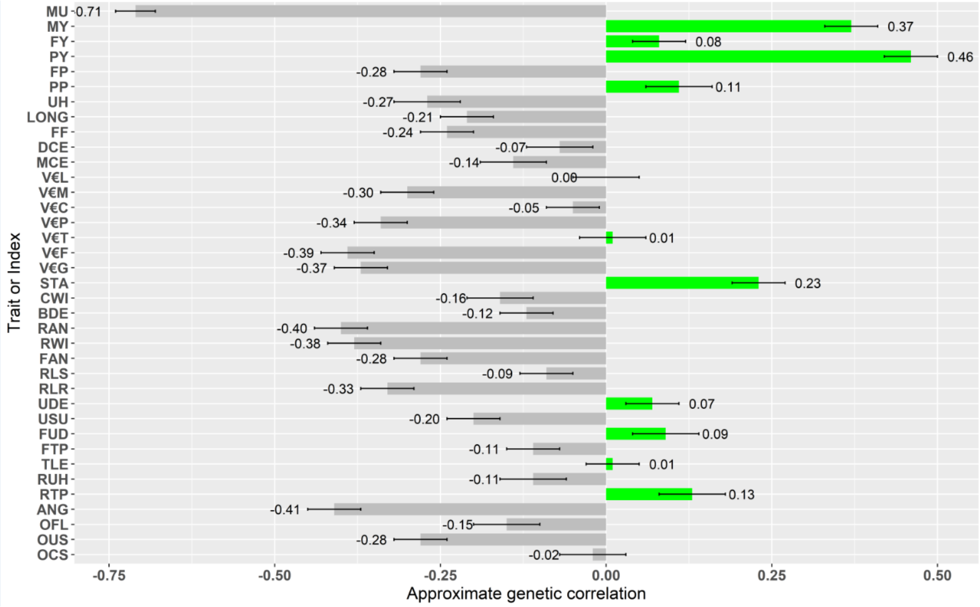


Figure 4-3. Approximate genetic correlations between nitrogen efficiency index (NEI) and other traits based on EBV of selected bulls (n = 736). Other traits included milk urea concentration (MU), milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP), protein percentage (PP), udder health (UH), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index (V€L), member economic index(V€M), capacity economic index (V€C), udder economic index (V€P), functional type economic index(V€T), functional economic index (V€F), global economic index (V€G), stature (STA), chest width (CWI), body depth (BDE), rump angle (RAN), rump width (RWI), foot angle (FAN), rear leg set (RLS), rear leg rear view (RLR), udder depth (UDE), udder support (USU), fore udder (FUD), front teat placement (FTP), teat length (TLE), rear udder height (RUH), rear teat placement (RTP), angularity (ANG), overall feet and leg score (OFL), overall udder score (OUS), and overall conformation score (OCS). (Note: black line is the estimated standard error obtained using the bootstrapped replicates method, and all standard errors of approximate genetic correlation < 0.05)

The genetic correlation found between NEI and production economic index (index combined of MY, FY, PY, FP, and PP) was close to zero. Unfavorable genetic correlation was observed between the NEI and member economic index (–0.30, representing leg and hoof health). Köck et al. (2018) reported positive genetic correlations between FE (ECM / DMI), energy efficiency (energy in milk/energy intake), and the incidence of lameness, which is consistent with the findings of this study. Genetic correlation found between NEI and capacity economic index (representing body size) was low (–0.05). An unfavorable genetic correlation was observed between the NEI and udder economic index (–0.34, representing the udder health), the reason for this correlation is the same as we suggested for UH. The functional type economic index is a combination of member economic, capacity economic, and udder economic indices. But surprisingly, the genetic correlation estimated between NEI and type economic index was close to zero (0.01). The genetic correlation between NEI and functional economic index was negative, probably because functional economic index is a combination of UH, LONG, FF, DCE, and MCE. The global economic index is a combination of 48% production economic index, 28% functional economic index, and 24% functional type economic index. Therefore, the NEI has a negative genetic correlation with global economic index. This means that the current in the Walloon Region of Belgium used global economic index will not lead to a favorable correlated response of NUE.

The genetic correlations estimated between the NEI and the conformation traits ranged from –0.41 to 0.23, which also showed a mutually confirming relationship with the above results. Like the UH results, negative genetic correlations were found between NEI and rear udder height and overall udder score, suggesting that increased NEI is not conducive to improving udder health. The NEI and angularity were negatively genetically correlated. This result shows that NEI can cause cows to become fat, which is also consistent with the finding that NEI is not beneficial to the FF. The estimated positive genetic correlation between NEI and stature (0.23) combined with the above results suggests that cows with a higher NUE may have become taller and fatter. The NEI, rump angle, and rump width were negatively genetically correlated, which was conducive to the relationships between NEI, FF, and LONG of cattle. However, the genetic correlation found between NEI and OCS was close to zero (–0.02), indicating that the increasing NEI would not affect conformation traits.

The genetic correlations between six traits used in the NEI composition and 37 traits (indices) are shown in Figure 4-1S to Figure 4-3S (appendix). The NINT had negative genetic correlations with MU, positive genetic correlations with production traits, functional traits, and global economic index, and its genetic correlations with OCS were close to zero. The genetic correlations of MTPN with MU were approximately 0, with production traits (except FP), functional traits (except UH and FF) and global economic index were positive, with OCS were negative. The MUNY had positive genetic correlations with MU, production traits (except FP and PP), functional traits (except FF), V€G, and OCS.

In summary, NEI is genetically positively correlated to production traits but negatively correlated to the health, function, indices (except production economic and functional type economic indices), and most conformation traits, consistent with our previous study of the predicted NUE (Chen et al., 2021c).

3.5. Potential Impact of the Selected NEI in Bull

The bottom-5% and top-5% NEI bull groups each had 37 bulls. The mean and SD for six traits (MY, PP, MU, NINT, MTPN, MUNY) of their daughters were shown in Table 4-4. The six traits of daughters were significantly different in the two bull groups (all $P < 0.01$). Compared to the bottom-5% bull group, the NINT and MUNY were lower, and MTPN was higher in the top-5% bull group. We got the same result even after correcting for the herd, test year, and parity effects of the six features. The results show that if we choose bulls with high NEI, their daughters may have the above characteristics.

Table 4-4. Mean and standard deviation (SD) of traits[1] for the daughters with bottom-5% and top-5% of nitrogen efficiency index sires

Group[2]	MY (kg/d)	PP (%)	MU (mg/dL)	NINT (kg/d)	MTPN (kg/d)	MUNY (kg/d x 1000)
Bottom-5% (n = 30,818)[3]						
Mean	31.30	3.18	24.20	0.48[4]	0.15	3.55
SD	8.21	0.36	8.60	0.08	0.04	1.59
Top-5% (n = 43,322)						
Mean	31.50	3.24	21.70	0.46	0.16	3.18
SD	8.45	0.37	8.12	0.08	0.04	1.48

[1] MY - Milk Yield; PP - Protein percentage; MU - Milk urea concentration; NINT - N intake; MTPN - Milk true protein N; MUNY – Milk urea N yield

[2] Group: P-value < 0.01 of the t-test between the two groups for all traits

[3] The 30,818 records in the bottom-5% group from 9,455 cows, and 43,322 records in the top-5% group from 13,506 cows

[4] The 7,059 records for NINT in the bottom-5% group from 2,633 cows, and 13,389 records in the top-5% group from 4,683 cows

Distributions between NEI of the selected bulls (n = 50) born from 2011 to 2014 and EBV of the eight selected traits (indices) are presented in Figure 4-4. Because the NEI has positive genetic correlations with MY and PY, we can easily find bulls with good NEI, MY, and PY at the same time (such as bull No. 45). NEI has negative genetic correlations with the remaining six traits (indices), so a small number of bulls fell in the upper right corner of the distribution map (better bulls). The good news is that a small number of bulls can have better EBV for most of the traits, including NEI, at the same time (such as Bull No. 15). Overall, 26% of 50 bulls had both higher NEI and V€G. This shows that using the NEI in genetic selection is feasible.

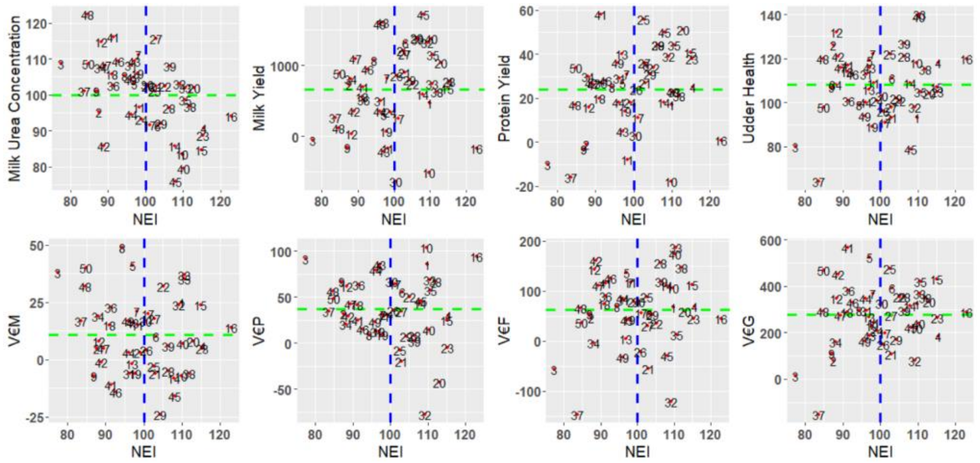


Figure 4-4. Distribution between standardized nitrogen efficiency index (NEI) of bulls (n = 50) born from 2011 to 2014 and EBV of other 8 selected traits or index (V€M - member economic index, V€P - udder economic index, V€F - functional economic index, V€G - global economic index) (PS: the dotted line represents the mean of the selection group)

The NEI defined in this study may need further optimization. Although we considered NINT, MTPN, and MUN (instead of urine and feces N), the N consumption for maintenance of body metabolism was not included, and the MUNY is only an indicator of nitrogen losses, with limited accuracy. Therefore, in further steps we plan to add live weight or body condition score to the existing NEI to improve it. The data of this study is based only on the first 50 DIM (limit from predicted NINT mode), which should be extended to the whole lactation. Currently, we only can use MU as an indicator for the whole lactation to indirectly increase the NUE and directly decrease the N pollution (Chen et al., 2021b). However, we should note that the genetic correlation between predicted NUE and MU was very low (Chen et al., 2021a) maybe because they are working in different processes of the NUE (Figure 4-1). The advantage of NEI is that it can consider both NUE and N pollution at the same time, and the genetic correlation was –0.71 between NEI and MU. In addition, dairy cows in early lactation are in a state of negative energy balance. Because the utilization of N by dairy cows is deeply dependent on energy availability, the genetic relationship between NEI and the energy status of dairy cows still needs to be explored.

4. Conclusions

The findings of this study showed low h^2 for traits of NINT, MTPN, and MUNY, ranging from 0.09 to 0.13. The genetic correlations found among NINT, MTPN, and MUNY were positive, except for that found between NINT and MUNY. The NEI defined based on NINT, MTPN and MUNY has a low h^2 (0.06) and moderate REL (mean 0.21), but according to the range of NEI values (–13.13 to 12.55 kg/d), it has genetic selection potential. The NEI showed positive genetic correlations with production traits (e.g., MY, PY), ranging from 0.08 to 0.46. The genetic correlations estimated between the NEI and MU, UH, LONG, FF, calving ease (DCE, MCE), and global economic index were negative, ranging from –071 to –0.07. Compared to the bottom-5% bull group, the NINT and MUNY were lower, and MTPN was higher in the top-5% bull group. The analyses of the selected bulls born from 2011 to 2014 showed that the bulls can have both higher NEI and global economic index than average at the same time. Finally, the developed NEI has the advantage of large-scale prediction and therefore has the potential to be routinely used in dairy cattle breeding in the future.

5. Acknowledgments

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Yansen Chen. Development of the used prediction equation was supported by the GplusE project which has received funding from the European Union’s Seventh Framework Programme (Brussels, Belgium) for research, technological development, and demonstration, under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. The authors acknowledge the support of the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE; Namur, Belgium) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech (Gembloux, Belgium) provided by the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the F National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grants No. T.0095.19 (PDR “DEEPSELECT”). The authors have not stated any conflicts of interest.

6. References

Blanchard, P. J., R. W. Everett, and S. R. Searle. 1983. Estimation of genetic trends and correlations for Jersey cattle. J. Dairy Sci. 66:1947–1954. https://doi.org/10.3168/jds.S0022 -0302(83)82033-5.

Bobbo, T., M. Penasa, A. Rossoni, and M. Cassandro. 2020. Short communication: Genetic aspects of milk urea nitrogen and new indicators of nitrogen efficiency in dairy cows. J. Dairy Sci. 103:9207–9212. https://doi.org/10.3168/jds.2020-18445.

Brito, L.F., H.R. Oliveira, K. Houlahan, P.A.S. Fonseca, S. Lam, A.M. Butty, D.J. Seymour, G. Vargas, T.C.S. Chud, F.F. Silva, C.F. Baes, A. Cánovas, F. Miglior, and F.S. Schenkel. 2020. Genetic mechanisms underlying feed utilization and implementation of genomic selection for improved feed efficiency in dairy cattle. Can. J. Anim. Sci. 100:587–604. https://doi.org/10.1139/cjas-2019-0193.

Calsamiglia, S., A. Ferret, C.K. Reynolds, N.B. Kristensen, and A.M. Van Vuuren. 2010. Strategies for optimizing nitrogen use by ruminants. Animal 4:1184–1196. https://doi.org/10.1017/S1751731110000911.

Cantalapiedra-Hijar, G., R.J. Dewhurst, L. Cheng, A.R.J. Cabrita, A.J.M. Fonseca, P. Nozière, D. Makowski, H. Fouillet, and I. Ortigues-Marty. 2018. Nitrogen iso topic fractionation as a biomarker for nitrogen use efficiency in ruminants: A meta-analysis. Animal 12:1827–1837. https://doi.org/10.1017/S1751731117003391.

Chen, Y., C. Grelet, S. Vanderick, and N. Gengler. 2021a. Relationship between pro xies of nitrogen use efficiency for dairy cows in early lactation. Accessed Nov ember 04, 2021. https://orbi.uliege.be/bitstream/2268/263087/1/EAAP21_71_YansenCHEN.pdf.

Chen, Y., H. Atashi, S. Vanderick, R.R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2021b. Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows. J. Dairy Sci. 104:12741–12755. https://doi.org/10.3168/jds.2021-20659.

Chen, Y., S. Vanderick, R.R. Mota, C. Grelet, G. Consortium, and N. Gengler. 2021c. Estimation of genetic parameters for predicted nitrogen use efficiency and losses in early lactation of Holstein cows. J. Dairy Sci. 104:4412–4423. https://doi.org/10.3168/jds.2020-18849.

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.

Jamrozik, J., G.J. Kistemaker, T.C.S. Chud, and L.R. Schaeffer. 2020. Estimates of genetic parameters for Holstein female fertility. Accessed June 14, 2021.

https://www.cdn.ca/Articles/GEBOCT2020/FE_VCE%20Report%20-%20September2020.pdf

Klaffenböck, M., A. Steinwidder, C. Fasching, G. Terler, L. Gruber, G. Mészáros, and J. Sölkner. 2017. The use of mid-infrared spectrometry to estimate the ration composition of lactating dairy cows. J. Dairy Sci. 100:5411–5421. https://doi.org/10.3168/jds.2016-12189.

Köck, A., M. Ledinek, L. Gruber, F. Steininger, B. Fuerst-Waltl, and C. Egger-Danner. 2018. Genetic analysis of efficiency traits in Austrian dairy cattle and their relationships with body condition score and lameness. J. Dairy Sci. 101:445–455. https://doi.org/10.3168/jds.2017-13281.

Li, B., P.M. VanRaden, E. Guduk, J.R. O'Connell, D.J. Null, E.E. Connor, M.J. VandeHaar, R.J. Tempelman, K.A. Weigel, and J.B. Cole. 2020. Genomic prediction of residual feed intake in US Holstein dairy cattle. J. Dairy Sci. 103:2477–2486. https://doi.org/10.3168/jds.2019-17332.

Li, B., W.F. Fikse, J. Lassen, M.H. Lidauer, P. Løvendahl, P. Mäntysaari, and B. Berglund. 2016. Genetic parameters for dry matter intake in primiparous Holstein, Nordic Red, and Jersey cows in the first half of lactation. J. Dairy Sci. 99:7232–7239. https://doi.org/10.3168/jds.2015-10669.

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.

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 May 06, 2021. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Parker Gaddis, K.L., P.M. VanRaden, R.J. Tempelman, K.A. Weigel, H.M. White, F. Peñagaricano, J.E. Koltes, J.E.P Santos, R.L. Baldwin, J.F. Burchard, J.W. Dürr, and M.J. VandeHaar. 2021. Implementation of feed saved evalutions in the U.S. Accessed May 05, 2021. https://www.uscdcb.com/wp-content/uploads/2021/05/FeedSaved_InterbullPresentation_v2_NoVid-Kristen.pdf.

Pryce, J.E., W.J. Wales, Y. De Haas, R.F. Veerkamp, and B.J. Hayes. 2014. Genomic selection for feed efficiency in dairy cattle. Animal 8:1–10. https://doi.org/10.1017/s1751731113001687.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed June 11, 2021. https://www.r-project.org/.

Spanghero, M., and Z.M. Kowalski. 2021. Updating analysis of nitrogen balance experiments in dairy cows. J. Dairy Sci. 104:7725–7737. https://doi.org/10.3168/jds.2020-19656.

Tempelman, R.J., D.M. Spurlock, M. Coffey, R.F. Veerkamp, L.E. Armentano, K.A. Weigel, Y. de Haas, C.R. Staples, E.E. Connor, Y. Lu, and M.J. VandeHaar. 2015. Heterogeneity in genetic and nongenetic variation and energy sink relationships for residual feed intake across research stations and countries. J. Dairy Sci. 98:2013–2026. https://doi.org/10.3168/jds.2014.8510.

Vallimont, J.E., C.D. Dechow, J.M. Daubert, M.W. Dekleva, J.W. Blum, C.M. Barlieb, W. Liu, G.A. Varga, A.J. Heinrichs, and C.R. Baumrucker. 2011. Short communication: Heritability of gross feed efficiency and associations with yield, intake, residual intake, body weight, and body condition score in 11 commercial Pennsylvania tie stalls. J. Dairy Sci. 94:2108–2113. https://doi.org/10.3168/jds.2010-3888.

Vallimont, J.E., C.D. Dechow, J.M. Daubert, M.W. Dekleva, J.W. Blum, W. Liu, G.A. Varga, A.J. Heinrichs, and C.R. Baumrucker. 2013. Short communication: Feed utilization and its associations with fertility and productive life in 11 commercial Pennsylvania tie-stall herds. J. Dairy Sci. 96:1251–1254. https://doi.org/10.3168/jds.2012-5712.

Vanderick, S., R. R. Mota, K. Wijnrocs, and N. Gengler. 2020. Description of the genetic evaluation systems used in the Walloon Region of Belgium. Accessed May 06, 2021. http://www.elinfo.be/docs/GESen2012.pdf.

VanRaden, P.M., J.B. Cole, and K.L. Parker Gaddis. 2018. Net merit as a measure of lifetime profit: 2018 revision. Accessed Aug. 03, 2021. https://www.aipl.arsusda.gov/reference/nmcalc-2018.htm.

WHO and FAO (World Health Organization and Food and Agriculture Organization of the United Nations). 2011. Codex Alimentarius: Milk and Milk Products. 2nd ed. WHO FAO, Rome, Italy.

Wisconsin Alumni Research Foundation. 2015. Milk Urea-N yield (MUNY) as a nutritional and environmental management tool for the dairy industry. US Pat. No. 9,955,673 B1.

7. Appendix

Table 4-1S. Reliability for bulls (n = 736) for the N intake in primiparous (NINT1), milk true protein N in primiparous (MTPN1), milk urea nitrogen yield in primiparous (MUNY1), N intake in multiparous (NINT2+), milk true protein N in multiparous (MTPN2+), and milk urea nitrogen yield in multiparous (MUNY2+)

Reliability	NINT1	MTPN1	MUNY1	NINT2+	MTPN2+	MUNY2+
Minimum	0.38	0.57	0.60	0.37	0.49	0.54
Maximum	0.94	0.98	0.99	0.96	0.99	0.99
Mean	0.63	0.82	0.85	0.67	0.84	0.84
Standard deviation	0.14	0.10	0.09	0.14	0.11	0.10

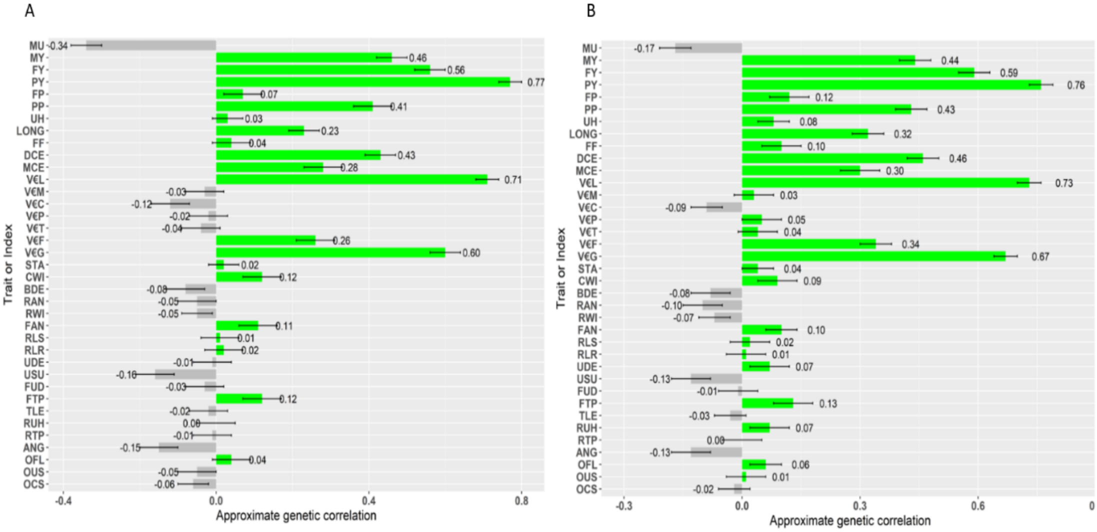


Figure 4-1S. Approximate genetic correlation between nitrogen intake (NINT) and other traits based on the EBV of selected bulls (n = 736) in primiparous (A, NINT1) and multiparous (B, NINT2+). Other traits included were milk urea concentration (MU), milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP), protein percentage (PP), udder health (UH), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index (V€L), member economic index(V€M), capacity economic index (V€C), udder economic index (V€P), functional type economic index(V€T), functional economic index (V€F), global economic index (V€G), stature (STA), chest width (CWI), body depth (BDE), rump angle (RAN), rump width (RWI), foot angle (FAN), rear leg set (RLS), rear leg rear view (RLR), udder depth (UDE), udder support (USU), fore udder (FUD), front teat placement (FTP), teat length (TLE), rear udder height (RUH), rear teat placement (RTP), angularity (ANG), overall feet and leg score (OFL), overall udder score (OUS), and overall conformation score (OCS). (Note: all standard errors of the approximate genetic correlations were < 0.05)

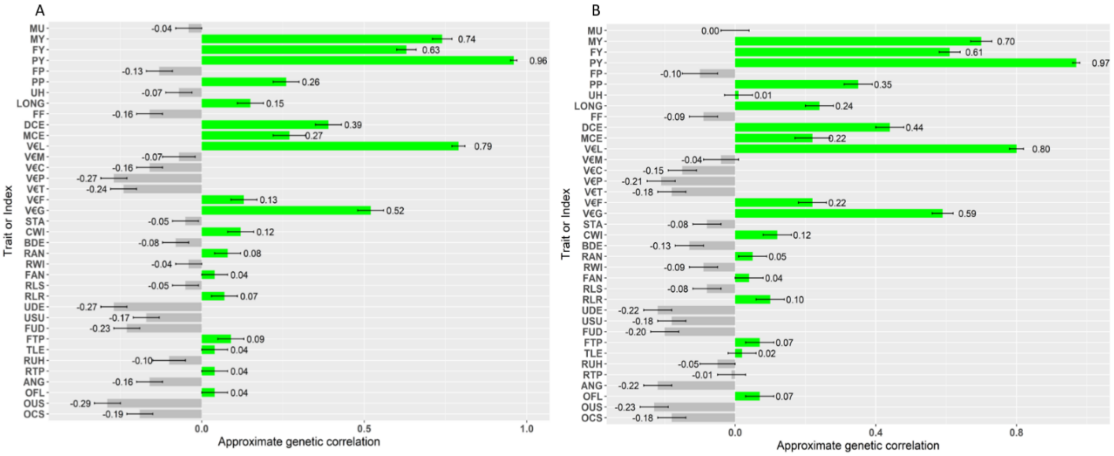


Figure 4-2S. Approximate genetic correlation between milk true protein nitrogen intake (MTPN) and other traits based on the EBV of selected bulls (n = 736) in primiparous (A, MTPN1) and multiparous (B, MTPN2+). Other traits included were milk urea concentration (MU), milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP), protein percentage (PP), udder health (UH), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index (V€L), member economic index(V€M), capacity economic index (V€C), udder economic index (V€P), functional type economic index(V€T), functional economic index (V€F), global economic index (V€G), stature (STA), chest width (CWI), body depth (BDE), rump angle (RAN), rump width (RWI), foot angle (FAN), rear leg set (RLS), rear leg rear view (RLR), udder depth (UDE), udder support (USU), fore udder (FUD), front teat placement (FTP), teat length (TLE), rear udder height (RUH), rear teat placement (RTP), angularity (ANG), overall feet and leg score (OFL), overall udder score (OUS), and overall conformation score (OCS). (Note: all stand errors of approximate genetic correlation were < 0.05)

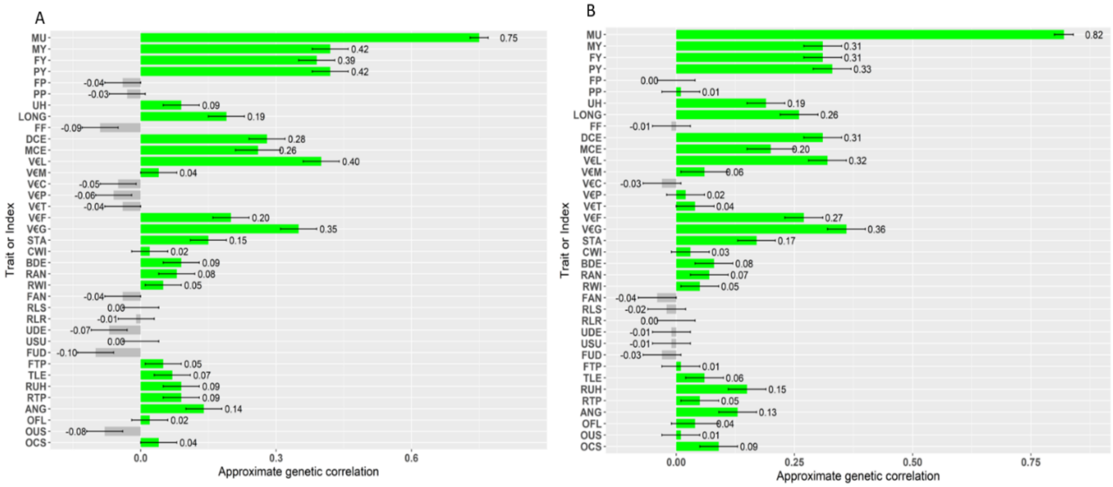


Figure 4-3S. Approximate genetic correlation between milk urea nitrogen yield (MUNY) and other traits based on the EBV of selected bulls (n = 736) in primiparous (A, MUNY1) and multiparous (B, MUNY2+). Other traits included were milk urea concentration (MU), milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP), protein percentage (PP), udder health (UH), longevity (LONG), female fertility (FF), direct calving ease (DCE), maternal calving ease (MCE), production economic index (V€L), member economic index(V€M), capacity economic index (V€C), udder economic index (V€P), functional type economic index(V€T), functional economic index (V€F), global economic index (V€G), stature (STA), chest width (CWI), body depth (BDE), rump angle (RAN), rump width (RWI), foot angle (FAN), rear leg set (RLS), rear leg rear view (RLR), udder depth (UDE), udder support (USU), fore udder (FUD), front teat placement (FTP), teat length (TLE), rear udder height (RUH), rear teat placement (RTP), angularity (ANG), overall feet and leg score (OFL), overall udder score (OUS), and overall conformation score (OCS). (Note: all stand errors of approximate genetic correlation were < 0.05)

5

Chapter V Validating genomic prediction for nitrogen efficiency index and its composition of Holstein cows in early lactation

Adapted from: Chen, Y., H. Atashi, R. R. Mota, C. Grelet, S. Vanderick, H., Hu, GplusE Consortium, and N. Gengler. Validating genomic prediction for nitrogen efficiency index and its composition of Holstein cows in early lactation. J. Dairy Sci. In review.

Foreword

The new nitrogen efficiency index (NEI) was built in Chapter 4. Genomic prediction is already used in animal breeding programs. This Chapter showed that NEI and its composition traits can be used in genomic prediction. The results provide evidence of the possibility for NEI to inform breeding strategies to improve farmers' profits and reduce nitrogen pollution.

Abstract

Nitrogen (N) use efficiency (NUE) is an economically important trait for dairy cows. Recently, we proposed a new N efficiency index (NEI), which simultaneously considers both NUE and N pollution. This study aimed to validate the genomic prediction for NEI and its composition traits, and to investigate the relationship between SNP effects estimated directly from NEI and indirectly from its composition traits. The composition of the NEI included genomic estimated breeding value (GEBV) of N intake (NINT), milk true protein N (MTPN), and milk urea N yield (MUNY). The edited data were 132,899 records on 52,064 cows distributed in 773 herds. The pedigree contained 122,368 animals. Genotypic data of 566,294 SNP was available for 4,514 individuals. In total, 4,413 cows (of them, 181 were genotyped) and 56 bulls (of them, 32 were genotyped) were selected as the validation populations. Linear regression method was used to validate the genomic prediction of NEI and its composition traits, and it was carried out by using BLUP and single-step genomic BLUP (ssGBLUP). The mean theoretical accuracies of validation populations from ssGBLUP were higher than that of BLUP for both NEI and its composition traits, ranging from 0.57 (MTPN) to 0.72 (NINT). The highest mean prediction accuracies for NEI and its composition traits were observed for the genotyped cows estimated under ssGBLUP, ranging from 0.48 (MTPN) to 0.66 (NINT). In addition, the SNP effects estimated from NEI composition traits multiple by relative weight were the same as the SNP effects estimated directly from NEI. This study preliminary showed that genomic prediction can be used for NEI. Moreover, the SNP effects of NEI can be indirectly calculated by using the SNP effects estimated from its composition traits. This study provided a basis for adding genomic information to establish and predict NEI as part of future routine genomic evaluation programs.

Key words: nitrogen use efficiency, SNP effect, mid-infrared spectra

1. Introduction

Livestock production contributes to 18% of the global human-induced greenhouse gases emissions in the form of carbon dioxide, methane, and nitrous oxide (Moran and Wall, 2011). The main sources of nitrous oxide production in cattle are deposited urine and manure (Rochette et al., 2014), which are mainly due to the protein content of the undigested feed. Protein is the most expensive component of dairy cows feed and its loss reduces farms' profits; therefore, improving the nitrogen (N) use efficiency (NUE) of dairy cows can potentially improve the profitability of dairy herds and reduce N pollution (Chen et al., 2021). To increase NUE in dairy cows, the management and breeding systems mainly focus on urinary N using milk urea concentration (MU) or N (MUN) (Kauffman and St-Pierre, 2001; Spek et al., 2013). Researchers in New Zealand, a highly grazing country, have recently investigated the potential consequences of using MUN in the genetic selection program to reduce N pollution in dairy cows (Correa-Luna et al., 2021; Marshall et al., 2021, 2022). The results of Marshall et al. (2021, 2022) showed that selection for low MUN breeding values can reduce urine N in cattle, whereas Correa-Luna et al. (2021) results showed otherwise. In the Walloon Region of Belgium, practically all cattle, to a certain degree, are grazing from April to September (Soyeurt et al., 2022). Farmers are looking for dairy cows that are most suitable for grazing to reduce production costs (Lefèvre et al., 2022). However, the oversupply of N for dairy cows in the grazing system may result in increased N pollution.

Recently, the results of a large international collaboration showed that predicting N excretion in dairy cows needs to include N intake (NINT) and MUN to more accurately predict N pollution, and more region-specific models are needed (Bougouin et al., 2022). We have proposed a new N efficiency index (NEI) combining the EBV of NINT, milk true protein N (MTPN), and MUN yield (MUNY) through the selection index, which considered both NUE and N pollution at the same time (Chen et al., 2022). The MUNY is proportional to urinary nitrogen excretion (Wisconsin Alumni Research Foundation, 2015). The NEI can keep NINT constant, increase MTPN, and decrease MUNY through genetic selection. However, the genomic information was not included when building the index. Genomic information is useful for selecting low-heritability traits, and has been widely used in animal breeding programs. The incorporation of non-genotyped animals into genome prediction is increasingly common through single-step genomic BLUP (ssGBLUP).

To the best of our knowledge, genomic prediction is currently performed only for traits but not for indices. The objects of this study were to validate the genomic prediction for the NEI and its composition traits, and to investigate the relationship between the single nucleotide polymorphism (SNP) effects obtained directly from NEI and indirectly from its composition traits. In this way, it will be possible to verify whether NEI can be directly genomically predicted and if this index reflects its composition at the SNP level.

2. Materials and methods

The study framework is shown in Figure 5-1.

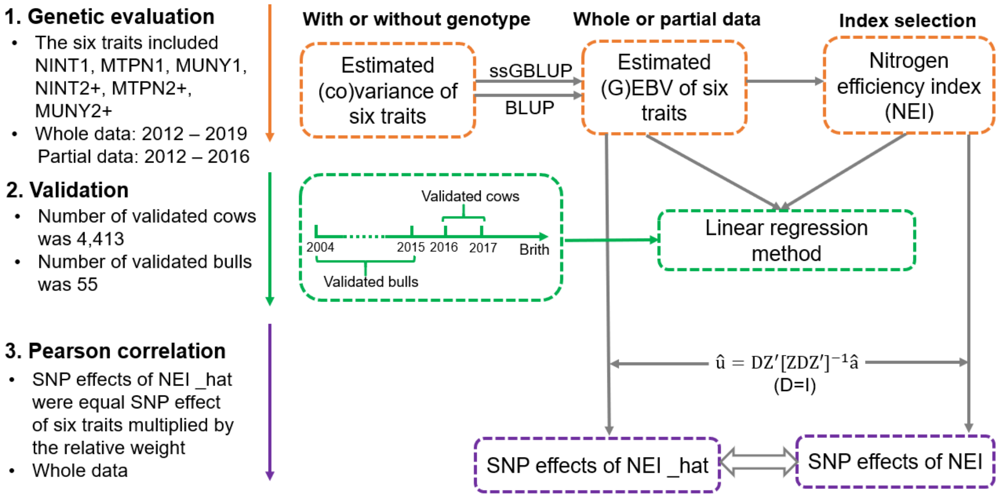


Figure 5-1. Workflow of validating genomic prediction for nitrogen efficiency index and its composition traits. NINT1 - the N intake in primiparous cows, MTPN1 - milk true protein N in primiparous cows, MUNY1 - milk urea N yield in primiparous cows, NINT2+ - N intake in multiparous cows, MTPN2+ - milk true protein N in multiparous cows, and MUNY2+ - milk urea N yield in multiparous cows.

2.1. Data

Phenotypic Data. As an exploratory study, the data between 2012 and 2019 were sampled from the official milk recording database in the Walloon Region of Belgium. The milk samples were analyzed by mid-infrared (MIR) spectrometry (commercial instruments from FOSS) to assess milk composition (including MU) and generate MIR spectra. All milk MIR spectra were standardized according to Grelet et al. (2015). The 3 features including NINT, MTPN, and MUNY were used in this study. The features were divided into 2 classes: primiparous and multiparous (lactations 2 to 5) and shown as NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+ (six traits in total). The formulas used to compute MTPN = [(Milk yield × Protein percent

/ 6.38) – MUNY] and MUNY = [(Milk urea concentration / 2.14) × Milk yield] (WHO and FAO, 2011). The NINT of each cow was predicted using the equation developed by Grelet et al. (2020). The determination coefficients and root mean square errors of validation for the NINT equation were 0.71 and 0.07 kg/d.

The data editing procedure used for all the included traits was the same as explained by Chen et al. (2022). Briefly, all records were restricted to the first 50 days in milk (DIM) because of the predicament of the NINT model. The NINT model was built based on data from the first 50 DIM only. For NINT, two criteria were considered in the filtering procedure: 1) the standardized Mahalanobis distance of the new predicted MIR spectra and the calibration dataset is ≤ 3; 2) the predicted value of NINT was restricted within the range of ± 3 standard deviations of the mean. Finally, 132,899 records, observed between 2012 and 2019 from 52,064 cows in 773 herds, were kept. The pedigree related to the dataset comprised 122,368 animals.

Genotypic data. the genotype data of 4,514 animals included in the pedigree were extracted from the cattle genetic evaluation system in the Walloon Region of Belgium. The used chip versions were BovineSNP50 K v1 to v3 (Illumina, San Diego, CA, USA). The SNPs common between all three chips were kept. Non-mapped SNPs, SNP located on sex chromosomes, and triallelic SNPs were excluded. A minimum GenCall Score of 0.15 and a minimum GenTrain Score of 0.55 were used to keep SNP (Wilmot et al., 2022). Then, the genotypes were imputed to HD using FImpute V2.2 software (Sargolzaei et al., 2014). The SNPs with Mendelian conflicts, and those with minor allele frequency less than 5% were excluded. The difference between observed heterozygosity and that expected under Hardy-Weinberg equilibrium was estimated, and if the difference was greater than 0.15, the SNP was excluded (Wiggans et al., 2009). Finally, 566,294 out of 730,539 SNPs, distributed on 29 Bos taurus autosomes, were kept.

2.2. (Co)variance Component Estimation

A six-trait (six traits; 3 features, 2 parity classes) repeatability model was used to estimate the variance components. All information on the model can be found in Chen et al. (2022). In brief, herd-year-season of calving, standardized DIM and its quadratic, and constant, linear, quadratic regression of standardized calving age (nested within parities) were used as fixed effects in this model, whereas non-genetic cow, non-genetic cow × parity (only for multiparous traits), additive animal genetic, and residual were used as random effects. However, to calculate the relationship between animals, the **H** or pedigree (**A**) relationship matrix was used. The **H** matrix combined **A** and genomic (**G**) based relationships matrices. The inverse of **H** as defined by Aguilar et al. (2010) is as follows:

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

where **A** is the numerator relationship matrix for all animals included in the pedigree; $\mathbf{A}_{22}$ is the numerator relationship matrix for genotyped animals; **G** is the genomic relationship matrix obtained using the function described by VanRaden (2008). In addition, the inverse of all matrices considers the coefficient of inbreeding between individuals (Lourenco et al., 2020).

Computations were performed using the BLUPF90 family of programs (Misztal et al., 2018). The (co)variance components for NINT, MTPN, and MUNY were estimated by Gibbs sampling using the models described by Chen et al. (2022). The posterior means of (co)variances, heritabilities (h^2), repeatabilities, genetic and phenotypic correlations were obtained using a single chain of 130,000 after a burn-in of 10,000 replicates. The formulas used to calculate the later four parameters have been previously described by Chen et al. (2021). The posterior convergence and approximated standard errors of (co)variances and four parameters were analyzed by POSTGIBBS90 (version 3.14). Repeatability was here referred to as within the 6 traits. For the traits in multiparous, repeatability was equal to the sum of the variances of the three random effects in the model (except for the residual) divided by the sum of the variances of the four random effects.

2.3. Estimated Breeding Values and Nitrogen Efficiency Index

By using the estimated genetic variance components, the EBV and genomic EBV (GEBV) were estimated for the six studied traits through BLUP and ssGBLUP, respectively. The (co)variance components were estimated for scenarios with (pedigree + phenotypes + genotypes) and without (pedigree + phenotypes) genomic data. The (G)EBVs of the six studied traits were estimated using the BLUPF90 program (version 1.71) through the same model used for estimating (co)variance components (Chen et al., 2022). The reliabilities (REL) of (G)EBVs for the studied traits based upon the whole data were calculated by the formula:

$$\mathrm{REL}_j = 1 - \frac{S\mathrm{E}_j^2}{\sigma_{uj}^2(1+f)}$$

where REL_j is the reliability of *j* trait, SE_j is the standard error of prediction for *j* trait retrieved from the inverse of the left-hand side matrix of the mixed model equations, σ_{uj}^2 is the additive genetic variance, and *f* is the mean inbreeding coefficient for all animals included in the pedigree.

The calculation methods of the NEI and its reliability were the same as those described by Chen et al. (2022). In brief, the NEI was defined as a combination of the (G)EBV of NINT, MTPN, and MUNY through the selection index. The index weights (**a**) of the six studied traits were calculated by selection responses, which assumed that the genetic selection responses for NINT, MTPN, and MUNY were 0, 1, and -1, respectively. This assumption allows genetic selection for NEI to keep NINT unchanged, increase MTPN, and decrease MUNY. The method for the REL of NEI from VanRaden et al. (2018):

$$\mathbf{REL_{NEI}} = \frac{\mathbf{w}\mathbf{G_C}\mathbf{w}'}{\mathbf{a}\mathbf{G_C}\mathbf{a}'}$$

where $\mathbf{REL_{NEI}}$ was the REL of NEI, $\mathbf{w}$ was obtained by multiplying the elements of $\mathbf{a}$ by the square root of reliability for the (G)EBV of the six studied traits, $\mathbf{G_C}$ was the genetic correlation matrix between the six studied traits.

2.4. Nitrogen Efficiency Index and Its Composition Traits Validation

The theoretical accuracies of NEI and its composition traits in the whole dataset were calculated first, which was considered the maximum accuracy of genetic selection in this dataset. The theoretical accuracy is equal to the square root of REL for NEI and each trait.

To verify the prediction accuracy of the genomic prediction in young animals, the linear regression method was used in this study (Legarra and Reverter, 2018). The basic step of the linear regression method is to calculate the evaluation metrics based on the breeding values from partial and whole datasets. Only the last three years of records were removed because the whole dataset was small (2012 to 2019, n = 132,899). Therefore, only the data from 2017 to 2019 was set to be missing and this reduced dataset was called the partial dataset (n = 93,993). Both variance components and breeding values need to be estimated again in the partial dataset.

The validation population (called the focal individuals) consisted of cows and bulls (Figure 5-1). Cows (n = 4,413) born after 2015 were used as the focal individuals (without phenotypes in the partial dataset), of which 181 cows were genotyped. We verified the cows with and without genotypic data, separately. In the case of genotyped cows, BLUP validation was also performed to show whether differences in results could be caused by different validation populations. Bulls (n = 55) were selected as the validation population under two criteria, of which 35 bulls were genotyped. The first criterion was bulls only have daughters born after 2015 with records of six traits. The second criterion was bulls with at least 5 daughters or at least

1 daughter and genotyped. These 55 bulls have 695 daughters with 872 records in this dataset. Those with and without genotypes were put together because of the small number of bulls that met both of the above criteria.

Four following metrics were used to measure prediction validation results in this study.

The population bias (μ_{wp}) was expected to be 0 if the evaluation was unbiased, which was defined as the following:

$$\mu_{wp} = \overline{\hat{u}_p} - \overline{\hat{u}_w}$$

where $\overline{\hat{u}_p}$ and $\overline{\hat{u}_w}$ were average (G)EBV of focal individuals in the partial and whole datasets, respectively.

The dispersion (b_{wp}) was expected to be 1 if the evaluation showed no dispersion, which was defined as the following:

$$b_{wp} = \frac{cov(\hat{u}_p,\ \hat{u}_w)}{var(\widehat{u_p}\,)}$$

where $\hat{u}_p$ and $\hat{u}_w$ were (G)EBV of focal individuals in the partial and whole datasets, respectively.

The prediction accuracy ($\widehat{acc}$) was expected to be 1 if the evaluation was perfect. It is the predictive accuracy of the breeding values of focal individuals in the partial dataset, which was defined as the following:

$$\widehat{acc} = \sqrt{\frac{cov(\hat{u}_p, \hat{u}_w)}{(1 - \overline{f})var(\hat{u}_p)}}$$

where $\overline{f}$ was the average inbreeding coefficient of the focal individuals in the partial and whole datasets, respectively. Other parameters were the same as described above.

The relative increased accuracy with added phenotype (inc_{phen}; Bermann et al., 2021) expected to low value (0) if the evaluations stability between the partial and whole datasets, which was defined as below:

$$inc_{phen} = \frac{\widehat{acc}_w - \widehat{acc}_p}{\widehat{acc}_p} = \frac{1}{cor(\hat{u}_p, \hat{u}_w)} - 1$$

where $\widehat{acc}_w$ and $\widehat{acc}_p$ were accuracy in the whole and partial datasets, respectively. Other elements were the same as described above.

2.5. Relationship between Nitrogen Efficiency Index and Its Composition Traits

The Pearson correlation coefficients between the direct SNP effects estimated from NEI and indirect SNP effects estimated from its composition traits were estimated. The SNP effect ($\hat{\mathbf{a}}$) for the NEI and its composition traits were estimated using the POSTGSF90 software (version 1.73; Aguilar et al., 2014). The formula for the ($\hat{\mathbf{a}}$) is the same as that described by Wang et al. (2012) but without iteration. The formula of $\hat{\mathbf{a}}$ was as follows:

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

where $\mathbf{D} = \mathbf{I}$, means that the weight for all SNPs is 1; $\mathbf{Z}_{\mathbf{g}}$ was an incidence matrix of genotype for each SNP; $\hat{\mathbf{u}}$ was a vector of GEBV of each trait for genotyped animals (n=4,514) in the whole dataset**.**

The SNP effects for the six studied traits were multiplied by **a** (relative weight) to calculate the SNP effects of NEI_hat. Then, the Pearson correlations between the SNP effects of NEI_hat and the SNP effects estimated directly from NEI were calculated.

3. Results and discussion

3.1. Genetic Parameter

The mean, h^2, and repeatability for the six studied traits are shown in Table 5-1. The NUE (MTPN divided by NINT) for primiparous (30.95%) was lower than that for multiparous (34.69%), which may be, at least in part, due to the more rumen development in multiparous cows. The average NUE values were within the normal range reported by previous studies (Cantalapiedra-Hijar et al., 2018; Spanghero and Kowalski, 2021). The h^2 and repeatability for the six traits ranged from 0.10 to 0.14, and 0.40 to 0.67, respectively. These results are similar to our previous results without using genotypic data (Chen et al., 2022).

Table 2-2 shows the genetic and phenotypic correlations among the six studied traits in this study. The same feature in primiparous and multiparous showed high genetic correlations, ranging from 0.85 to 0.96. The NINT showed medium positive genetic correlations with MTPN (0.45 to 0.53), but low negative genetic correlations with MUNY (−0.08 to −0.20). These results suggest that NINT may have a greater genetic influence on MTPN, but less on MUNY. Medium positive genetic correlations were found between MTPN and MUNY (0.39 to 0.56). Compared to genetic correlations, phenotypic correlations were relatively lower. In general, genetic and phenotypic correlations among the six traits are consistent with our previous study (Chen et al., 2022). However, lower genetic correlations were estimated between NINT and MUNY when using genotypic data.

Table 5-1. Description statistics of the six studied traits[1] and their heritabilities and repeatabilities were estimated by single-step genomic BLUP in the whole dataset

Parity class[2]	NINT (kg/d)	MTPN (kg/d)	MUNY (g/d)
1 (n = 40,916)			
Mean	0.42	0.13	2.93
SD[3]	0.06	0.03	1.08
Heritability	0.14 ± 0.01	0.13 ± 0.01	0.14 ± 0.01
Repeatability	0.45 ± 0.01	0.61 ± 0.00	0.40 ± 0.01
2+ (n = 91,938)			
Mean	0.49	0.17	3.74
SD	0.07	0.04	1.56
Heritability	0.13 ± 0.01	0.11 ± 0.01	0.10 ± 0.01
Repeatability	0.53 ± 0.00	0.67 ± 0.00	0.43 ± 0.00

[1] NINT - N intake; MTPN – Milk true protein N; MUNY – Milk urea N yield

[2] Parity class: Based on the parity, the data were divided into 2 classes including primiparous (class 1) and multiparous (class 2+)

[3] SD: standard deviation

Table 5-2. Genetic correlations (above the diagonal) and phenotypic correlations (below the diagonal) between the six studied traits were estimated by single-step genomic BLUP in the whole dataset

Traits[1]	NINT1	MTPN1	MUNY1	NINT2+	MTPN2+	MUNY2+
NINT1		0.48 ± 0.04	-0.20 ± 0.05	0.87 ± 0.02	0.50 ± 0.04	-0.16 ± 0.05
MTPN1	0.36 ± 0.01		0.56 ± 0.04	0.53 ± 0.04	0.85 ± 0.03	0.46 ± 0.05
MUNY1	0.06 ± 0.01	0.37 ± 0.00		-0.03 ± 0.04	0.39 ± 0.04	0.89 ± 0.03
NINT2+	0.17 ± 0.01	0.14 ± 0.01	0.02 ± 0.01		0.61 ± 0.03	-0.08 ± 0.05
MTPN2+	0.11 ± 0.01	0.20 ± 0.01	0.11 ± 0.01	0.43 ± 0.00		0.45 ± 0.04
MUNY2+	-0.02 ± 0.01	0.09 ± 0.01	0.16 ± 0.01	0.14 ± 0.00	0.41 ± 0.00	

[1]: NINT1 - N intake in primiparous cows; MTPN1 – Milk true protein N in primiparous cows; MUNY1 - Milk urea nitrogen yield in primiparous cows; NINT2+ - N intake in multiparous cows; MTPN2+ - Milk true protein N in multiparous cows; MUNY2+ - Milk urea nitrogen yield in multiparous cows

Table 5-2 shows the genetic and phenotypic correlations among the six studied traits in this study. The same feature in primiparous and multiparous showed high genetic correlations, ranging from 0.85 to 0.96. The NINT showed medium positive genetic correlations with MTPN (0.45 to 0.53), but low negative genetic correlations with MUNY (−0.08 to −0.20). These results suggest that NINT may have a greater genetic influence on MTPN, but less on MUNY. Medium positive genetic correlations were found between MTPN and MUNY (0.39 to 0.56). Compared to genetic correlations, phenotypic correlations were relatively lower. In general, genetic and phenotypic correlations among the six traits are consistent with our previous study (Chen et al., 2022). However, lower genetic correlations were estimated between NINT and MUNY when using genotypic data.

3.2. Validated the Nitrogen Efficiency Index and Its Composition Traits

The mean theoretical accuracies of focal individuals from ssGBLUP were higher than that of BLUP for both NEI and its composition traits (Table 5-3), which is consistent with Cesarani et al. (2021). The theoretical accuracy from ssGBLUP is higher because it has a small prediction error.

Table 5-3. The mean theoretical accuracies in the whole dataset for genotype cows, un-genotype cows, and bulls of nitrogen efficiency index (NEI) and its composition traits[1]

	NINT1	MTPN1	MUNY1	NINT2+	MTPN2+	MUNY2+	NEI
Genotype cows (n=181)[2]							
BLUP	0.63	0.58	0.61	0.61	0.57	0.59	0.57
ssGBLUP[3]	0.72	0.67	0.69	0.71	0.67	0.69	0.66
Un-genotype cows (n=4,232)							
BLUP	0.62	0.58	0.60	0.60	0.57	0.59	0.57
ssGBLUP	0.63	0.59	0.61	0.61	0.57	0.59	0.58
Bulls (n=55)[4]							
BLUP	0.64	0.60	0.62	0.61	0.58	0.60	0.59
ssGBLUP	0.70	0.66	0.68	0.68	0.64	0.67	0.65

[1]: NINT1 - N intake in primiparous cows; MTPN1 – Milk true protein N in primiparous cows; MUNY1 - Milk urea N yield in primiparous cows; NINT2+ - N intake in multiparous cows; MTPN2+ - Milk true protein N in multiparous cows; MUNY2+ - Milk urea N yield in multiparous cows

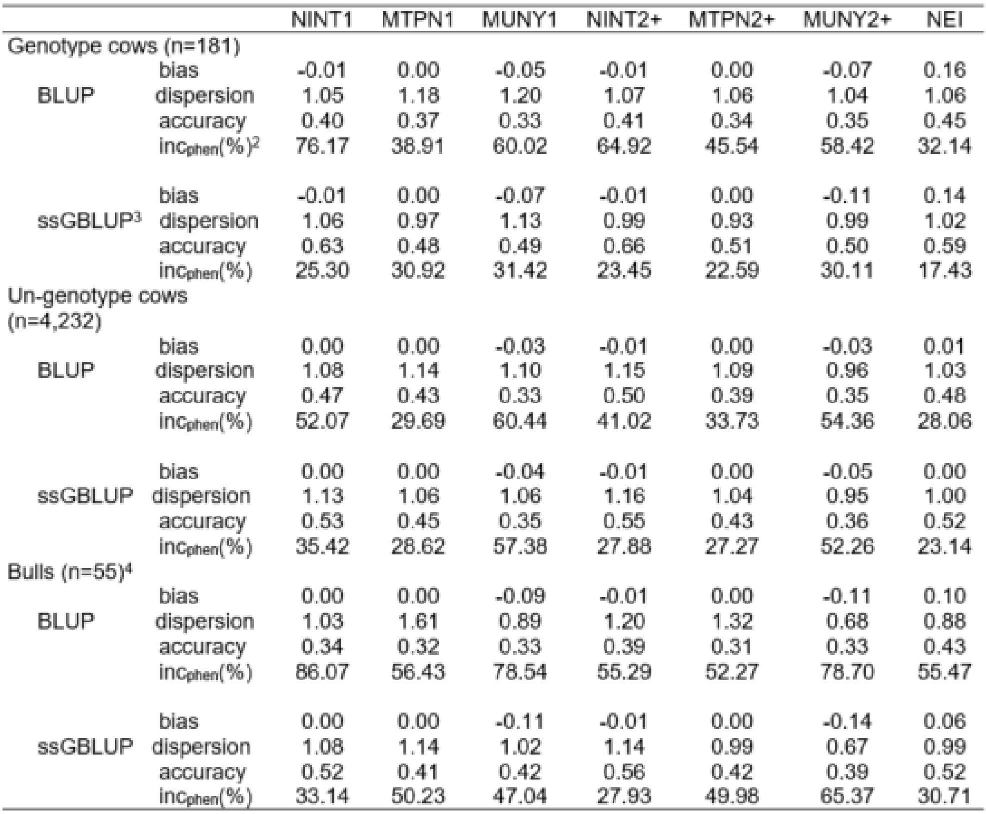
[2]: Ranges of standard deviation for genotype cows, un-genotype cows, and bulls are 0.06 to 0.10, 0.05 to 0.06, and 0.03 to 0.05, respectively

[3]: ssGBLUP - Single-step genomic BLUP

[4]: 35 of 55 bulls had genotype

The validated results of genomic prediction for NEI and its composition traits in genotyped cows, non-genotyped cows, and bulls through BLUP and ssGBLUP are presented in Table 5-4.

Table 5-4. Validated nitrogen efficiency index (NEI) and its composition traits[1] by linear regression



		NINT1	MTPN1	MUNY1	NINT2+	MTPN2+	MUNY2+	NEI
Genotype cows (n=181)								
	bias	-0.01	0.00	-0.05	-0.01	0.00	-0.07	0.16
BLUP	dispersion	1.05	1.18	1.20	1.07	1.06	1.04	1.06
	accuracy	0.40	0.37	0.33	0.41	0.34	0.35	0.45
	inc$_{phen}$(%)[2]	76.17	38.91	60.02	64.92	45.54	58.42	32.14
	bias	-0.01	0.00	-0.07	-0.01	0.00	-0.11	0.14
ssGBLUP[3]	dispersion	1.06	0.97	1.13	0.99	0.93	0.99	1.02
	accuracy	0.63	0.48	0.49	0.66	0.51	0.50	0.59
	inc$_{phen}$(%)	25.30	30.92	31.42	23.45	22.59	30.11	17.43
Un-genotype cows (n=4,232)								
	bias	0.00	0.00	-0.03	-0.01	0.00	-0.03	0.01
BLUP	dispersion	1.08	1.14	1.10	1.15	1.09	0.96	1.03
	accuracy	0.47	0.43	0.33	0.50	0.39	0.35	0.48
	inc$_{phen}$(%)	52.07	29.69	60.44	41.02	33.73	54.36	28.06
	bias	0.00	0.00	-0.04	-0.01	0.00	-0.05	0.00
ssGBLUP	dispersion	1.13	1.06	1.06	1.16	1.04	0.95	1.00
	accuracy	0.53	0.45	0.35	0.55	0.43	0.36	0.52
	inc$_{phen}$(%)	35.42	28.62	57.38	27.88	27.27	52.26	23.14
Bulls (n=55)[4]								
	bias	0.00	0.00	-0.09	-0.01	0.00	-0.11	0.10
BLUP	dispersion	1.03	1.61	0.89	1.20	1.32	0.68	0.88
	accuracy	0.34	0.32	0.33	0.39	0.31	0.33	0.43
	inc$_{phen}$(%)	86.07	56.43	78.54	55.29	52.27	78.70	55.47
	bias	0.00	0.00	-0.11	-0.01	0.00	-0.14	0.06
ssGBLUP	dispersion	1.08	1.14	1.02	1.14	0.99	0.67	0.99
	accuracy	0.52	0.41	0.42	0.56	0.42	0.39	0.52
	inc$_{phen}$(%)	33.14	50.23	47.04	27.93	49.98	65.37	30.71

[1]: NINT1 - N intake in primiparous cows; MTPN1 – Milk true protein N in primiparous cows; MUNY1 - Milk urea nitrogen yield in primiparous cows; NINT2+ - N intake in multiparous cows; MTPN2+ - Milk true protein N in multiparous cows; MUNY2+ - Milk urea nitrogen yield in multiparous cows

[2]: inc$_{phen}$(%) - increase reliability when adding phenotypic

[3]: ssGBLUP - single-step genomic BLUP

First, the results of BLUP and ssGBLUP were compared. The estimated μ_{wp} (nearly 0) and b_{wp}(nearly 1) for the NEI and its composition traits of BLUP and ssGBLUP in genotyped cows indicate that bias and dispersion for genomic prediction between the two methods are similar. Similar results were found for non-genotyped cows and bulls, except the b_{wp} of bulls. Gao et al. (2021) observed similar results when comparing

BLUP and ssGBLUP, even though they used more genotyped animals than non-genotyped animals. As expected, the $\widehat{acc}$ of the NEI and its composition traits estimated by ssGBLUP in genotyped cows were higher than that estimated by BLUP, ranging from 0.48 to 0.66. The **H** relationship matrix, used in the ssGBLUP, may explain this finding (Cesarani et al., 2021). Even for non-genotyped animals, the $\widehat{acc}$ of the NEI and its composition traits estimated by ssGBLUP were slightly higher than that estimated by BLUP, ranging from 0.36 to 0.55. The genotyped individuals are relatives to a part of non-genotyped animals which can partly explain this finding. The $\widehat{acc}$ of the NEI and its composition traits estimated by ssGBLUP in bulls were higher than that estimated by BLUP, ranging from 0.39 to 0.56. Similarly, multiple studies have shown that the accuracy of ssGBLUP is higher than BLUP (Bermann et al., 2021; Cesarani et al., 2021). The inc_{phen} of the NEI and its composition traits in genotyped cows and bulls estimated by ssGBLUP were lower than that estimated by BLUP, which is in line with Bermann et al. (2021). However, the inc_{phen} of NEI and its composition traits in non-genotyped animals estimated by ssGBLUP and BLUP were similar, except for NINT. This suggests that when making genomic predictions for non-phenotyped cows, providing their own genotypic data is a good predictor.

On the other hand, the obtained results were compared between genotyped cows, non-genotyped cows, and bulls. The μ_{wp} of NINT and MTPN estimated by ssGBLUP (or BLUP) for genotyped cows, non-genotyped cows, and bulls were similar, while the μ_{wp} of MUNY and NEI estimated by ssGBLUP (or BLUP) for genotyped cows and bulls were higher than that estimated for non-genotyped cows. The b_{wp} of NEI and its composition traits estimated by ssGBLUP (or BLUP) were similar between genotyped and non-genotyped cows. However, the b_{wp} of NEI and its composition traits estimated by ssGBLUP (or BLUP) in bulls were worse than the cows. The $\widehat{acc}$ of the NEI and its composition traits estimated by ssGBLUP in genotyped cows were higher than that in non-genotyped cows and bulls, which can be a consequence of using extra information available (SNP data) for genotyped cows. This is an advantage of using genomic selection for low heritability traits (Parker Gaddis et al., 2014). However, the $\widehat{acc}$ of the NEI and its composition traits estimated by BLUP in genotyped cows and bulls (already remove genotypic data) were lower than that in non-genotyped cows, which can be attributed to the smaller number of cows in the genotyped group (n = 181) and bulls (n = 55) compared with non-genotyped (n = 4,232). The inc_{phen} of NEI and its composition traits estimated by ssGBLUP (or BLUP) ranged from 17.43% (32.14%) to 31.42% (76.17%) in genotyped cows, from 23.14% (28.06%) to 57.38% (60.44%) in non-genotyped cows, and from 27.93% (52.27%) to 65.73% (86.07%) in bulls.

In short, these findings showed that genomic information benefits genomic predictions for NEI and its composition traits. The highest mean prediction accuracies for NEI and its composition traits were estimated for genotyped cows using ssGBLUP. Furthermore, we need to mention that even for non-genotyped and non-phenotype animals in the partial dataset, the prediction accuracy of genomic prediction for NEI was nearly theoretical accuracy by ssGBLUP (0.52 vs. 0.58).

3.3. Relationship between Nitrogen Efficiency Index and Its Composition Traits

Figure 5-2 shows the 566,294 SNPs effects values of the NEI, ranging from -0.003 to 0.003. The Pearson correlation between SNP effects estimated directly from NEI and indirectly from its composition traits was 1, and the intercept was 0 (Figure 5-2), which indicates that indirect genomic prediction can be performed on the NEI index.

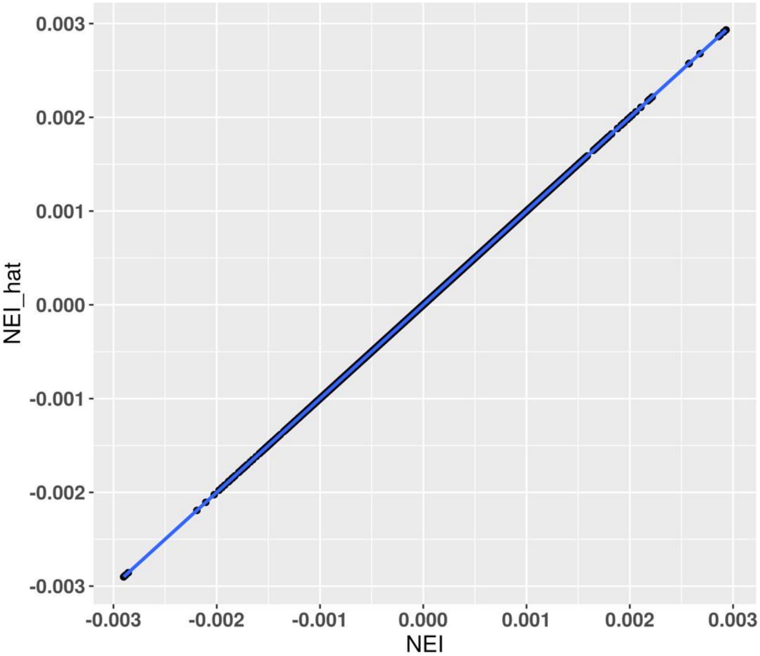


Figure 5-2. Relationship between SNP effects is estimated directly from the nitrogen efficiency index (NEI) and indirectly from its composition traits (NEI-hat).

Over the past few decades, the breeding goal of dairy cattle has gradually shifted from increasing milk production to balanced breeding (Miglior et al., 2017), which has resulted in more traits being added to the breeding system. With increasing the number of traits in a breeding program, breeders usually classify the traits to generate an index, and then combine them to generate a total index for breeding. Our results showed that the SNP effects estimated directly from the generated index are completely consistent with those estimated from its composition traits (Figure 5-2). Similarly, if the SNP effects of the total index are obtained, it is more convenient to perform genomic prediction on young bulls with unknown traits. Indirect genome prediction can also reduce the computational cost of large-scale genomic selection (Tsuruta et al., 2021), facilitating the implementation of genomic prediction at the national level.

Another tentative advantage of indirect genomic prediction of NEI is the potential reduction in prediction error. There are potential errors in the calculation of breeding values for each of the six traits, which may lead to a reduction in the accuracy of NEI. The potential improvement in the accuracy of the NEI can be achieved by reducing the number of steps in its calculation. However, the theoretical accuracy of indirect genomics predictions of NEI needs review, as Garcia et al. (2022) did for post-weaning weight gain. Furthermore, the effectiveness of indirect genomic prediction for indices needs to be validated on more indices.

3.4. Limitations and Perspectives

The results of this study also require several points of attention. First, the NEI is a proxy for the NUE of dairy cattle, and therefore not identical to the genetic merit of direct NUE. Moreover, there were some prediction errors in the NEI calculation process, such as the prediction models of the MIR-based phenotypes and the estimation of relevant GEBV. Two approaches (Data section) were adopted in this study to avoid these problems. Additional reference phenotypes are expected to be added, improving the prediction models of the MIR-based phenotypes. Furthermore, Tiplady et al. (2022) recently reported that genetic correlations between directly measured, and MIR-predicted fatty acid and protein fractions were typically high. This suggests that MIR-based phenotypes potentially represent still valid estimation of genetic contributions to directly measured traits. The estimation of relevant GEBV will also become more accurate with an increasing number of genotyped animals.

This study should be considered an exploration study, also based on a limited dataset. This implies that before constructing a national-level NEI combining genomic information and its integration into a global breeding goal additional work will be required. Although the mean prediction accuracies (0.52-0.59) of NEI in focal individuals from ssGBLUP are not very high, they are not too far from mean theoretical accuracies (0.58-0.66). Our results show that the average reliability (square of theoretical accuracy) of the bulls (n = 736) was 0.68 when establishing NEI (without genomic information) (Chen et al., 2022). This suggests that prediction accuracy for NEI can be improved with the increased number of records.

The small number of records extracted from the official milk recording database is the main reason for the low accuracy obtained. Especially, NINT data used was still limited to the first 50 DIM because the currently available MIR prediction equation is established using only data from the first 50 DIM. However, extending it to the entire lactation period is planned when relevant reference data will be available. This will be our focus as the next step, which should also increase the reliability of NEI because we could use MIR-predicted phenotypes across the whole lactation. Another possible way to increase the reliability of NEI is to increase the number of genotyped animals. The number of cows genotyped has increased rapidly in recent years (Wiggans and Carrillo, 2022), and this is not only the case in the USA but also, to a lesser degree, in the Walloon Region of Belgium.

The routine use of NEI would start with its integration into a broader breeding goal. In the context of the Walloon genetic evaluation system, which defines sub-indexes (Vanderick et al., 2022), it would be relevant to use the NEI as an N efficiency sub-index or as part of a broader efficiency sub-index. Breeding for improved NUE in cattle could be initiated by this approach.

4. Conclusions

This study showed that genomic prediction can be used for NEI. Furthermore, even though animals in the partial dataset were without genotypic and phenotypic data, the mean prediction accuracy of genomic prediction for NEI was nearly the mean theoretical accuracy by ssGBLUP (0.52 vs. 0.58). Simultaneously, we considered the feasibility of the direct and indirect genomic prediction for the NEI index, which may be more beneficial to the computation of larger datasets at the national level. This study provided a basis for adding genomic information to establish and predict NEI as part of future routine genomic evaluation programs.

5. Acknowledgments

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Yansen Chen. The authors would like to thank the European Union’s Seventh Framework Program for funding the GplusE project under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. The authors acknowledge the support of the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et

Environnement, SPW-DGARNE; Namur, Belgium) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech (Gembloux, Belgium) provided by the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the F National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grants No. T.0095.19 (PDR “DEEPSELECT”) and J.0174.18 (CDR “PREDICT-2”). The authors have not stated any conflicts of interest.

6. Code availability

The R code for calculating the NEI and its reliability with example data can be found on GitHub (https://github.com/Yansen0515/GP_NEI).

7. 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. Proceedings of the 10th world congress of genetics applied to livestock production, Vancouver, BC, Canada. Accessed August 21, 2022. http://www.ainfo.inia.uy/digital/bitstream/item/15445/1/Aguilar-et-al.-2014.-WCGALP.pdf.

Bermann, M., A. Legarra, M.K. Hollifield, Y. Masuda, D. Lourenco, and I. Misztal. 2021. Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: An application in chicken mortality. J. Anim. Breed. Genet. 138:4–13. https://doi.org/10.1111/jbg.12507.

Bougouin, A., A. Hristov, J. Dijkstra, M.J. Aguerre, S. Ahvenjärvi, C. Arndt, A. Bannink, A.R. Bayat, C. Benchaar, T. Boland, W.E. Brown, L.A. Crompton, F. Dehareng, I. Dufrasne, M. Eugène, E. Froidmont, S. van Gastelen, P.C. Garnsworthy, A. Halmemies-Beauchet-Filleau, S. Herremans, P. Huhtanen, M. Johansen, A. Kidane, M. Kreuzer, B. Kuhla, F. Lessire, P. Lund, E.M.K. Minnée, C. Muñoz, M. Niu, P. Nozière, D. Pacheco, E. Prestløkken, C.K. Reynolds, A. Schwarm, J.W. Spek, M. Terranova, A. Vanhatalo, M.A. Wattiaux, M.R. Weisbjerg, D.R. Yáñez-Ruiz, Z. Yu, and E. Kebreab. 2022. Prediction of nitrogen excretion from data on dairy cows fed a wide range of diets compiled in an intercontinental database: A meta-analysis. J. Dairy Sci. 105:7462–7481. https://doi.org/10.3168/jds.2021-20885.

Cantalapiedra-Hijar, G., R.J. Dewhurst, L. Cheng, A.R.J. Cabrita, A.J.M. Fonseca, P. Nozière, D. Makowski, H. Fouillet, and I. Ortigues-Marty. 2018. Nitrogen isotopic fractionation as a biomarker for nitrogen use efficiency in ruminants:

A meta-analysis. Animal 12:1827–1837. https://doi.org/10.1017/S1751731117003391.

Cesarani, A., A. Garcia, J. Hidalgo, L. Degano, D. Vicario, N.P.P. Macciotta, and D. Lourenco. 2021. Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. J. Dairy Sci. 104:5719–5727. https://doi.org/10.3168/jds.2020-19838.

Chen, Y., H. Atashi, C. Grelet, S. Vanderick, H. Hu, and N. Gengler. 2022. Defining a nitrogen efficiency index in Holstein cows and assessing its potential impact on the breeding program of bulls. J. Dairy Sci. 105: 7575–7587. https://doi.org/10.3168/jds.2021-21681.

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. 104:4412–4423. https://doi.org/10.3168/jds.2020-18849.

Correa-Luna, M., D. Donaghy, P. Kemp, M. Schutz, and N. López-Villalobos. 2021. Nitrogen use efficiency and excretion in grazing cows with high and low milk urea nitrogen breeding values. Sustainability 13:1–12. https://doi.org/10.3390/su13179827.

Gao, H., G. Su, J. Jensen, P. Madsen, O.F. Christensen, B. Ask, B.G. Poulsen, T. Ostersen, and B. Nielsen. 2021. Genetic parameters and genomic prediction for feed intake recorded at the group and individual level in different production systems for growing pigs. Genet. Sel. Evol. 53:1–12. https://doi.org/10.1186/s12711-021-00624-3.

Garcia, A., I. Aguilar, A. Legarra, S. Tsuruta, I. Misztal, and D. Lourenco. 2022. Theoretical accuracy for indirect predictions based on SNP effects from single-step GBLUP. Genet. Sel. Evol. 54:1–11. https://doi.org/10.1186/s12711-022-00752-4.

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., 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.

Kauffman, A.J., and N.R. St-Pierre. 2001. The relationship of milk urea nitrogen to urine nitrogen excretion in Holstein and Jersey cows. J. Dairy Sci. 84:2284–2294. https://doi.org/10.3168/jds.S0022-0302(01)74675-9.

Lefèvre, A., V. Decruyenaere, and E. Froidmont. 2022. Transition towards more robust cows on Walloon dairy farms for a better use of forage resources. Dair'Innov 2022 congress, Namur, Belgium.

Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. Genet. Sel. Evol. 50:1–18. https://doi.org/10.1186/s12711-018-0426-6.

Lourenco, D., A. Legarra, S. Tsuruta, Y. Masuda, I. Aguilar, and I. Misztal. 2020. Single-step genomic evaluations from theory to practice: using SNP chips and sequence data in blupf90. Genes (Basel). 11:1–32. https://doi.org/10.3390/genes11070790.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2021. Nitrogen balance of dairy cows divergent for milk urea nitrogen consuming either plantain or perennial ryegrass. Animals 11:1–18. https://doi.org/10.3390/ani11082464.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2022. Urine and fecal excretion patterns of dairy cows divergent for milk urea nitrogen breeding values consuming either a plantain or ryegrass diet. J. Dairy Sci. 105:4218–4236. https://doi.org/10.3168/jds.2021-21490.

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. A. L. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. Manual for BLUPF90 family programs. University of Georgia. Accessed Aug. 01, 2022. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Moran, D., and E. Wall. 2011. Livestock production and greenhouse gas emissions: Defining the problem and specifying solutions. Anim. Front. 1:19–25. https://doi.org/10.2527/af.2011-0012.

Parker Gaddis, K.L., J.B. Cole, J.S. Clay, and C. Maltecca. 2014. Genomic selection for producer-recorded health event data in US dairy cattle. J. Dairy Sci. 97:3190–3199. https://doi.org/10.3168/jds.2013-7543.

Rochette, P., M.H. Chantigny, N. Ziadi, D.A. Angers, G. Bélanger, É. Charbonneau, D. Pellerin, C. Liang, and N. Bertrand. 2014. Soil nitrous oxide emissions after deposition of dairy cow excreta in eastern Canada. J. Environ. Qual. 43:829–841. https://doi.org/10.2134/jeq2013.11.0474.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genom. 15:1–12. https://doi.org/10.1186/1471-2164-15-478.

Soyeurt, H., C. Gerards, C. Nickmilder, J. Bindelle, S. Franceschini, F. Dehareng, D. Veselko, C. Bertozzi, N. Gengler, A. Marvuglia, A. Bayram, and A. Tedde. 2022. Prediction of indirect indicators of a grass-based diet by milk Fourier Transform mid-infrared spectroscopy to assess the feeding typologies of dairy farms. Animals 12:1–13. https://doi.org/10.3390/ani12192663.

Spanghero, M., and Z.M. Kowalski. 2021. Updating analysis of nitrogen balance experiments in dairy cows. J. Dairy Sci. 104:7725–7737. https://doi.org/10.3168/jds.2020-19656.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, and A. Bannink. 2013. A review of factors influencing milk urea concentration and its relationship with urinary urea excretion in lactating dairy cattle. J. Agric. Sci. 151:407–423. https://doi.org/10.1017/S0021859612000561.

Tiplady, K.M., T.J. Lopdell, R.G. Sherlock, T.J.J. Johnson, R.J. Spelman, B.L. Harris, S.R. Davis, M.D. Littlejohn, and D.J. Garrick. 2022. Comparison of the genetic characteristics of directly measured and Fourier-transform mid-infrared-predicted bovine milk fatty acids and proteins. J. Dairy Sci. In press. https://doi.org/10.3168/jds.2022-22089.

Tsuruta, S., D.A.L. Lourenco, Y. Masuda, T.J. Lawlor, and I. Misztal. 2021. Reducing computational cost of large-scale genomic evaluation by using indirect genomic prediction. JDS Commun. 2:356–360. https://doi.org/10.3168/jdsc.2021-0097.

Vanderick, S., R. R. Mota, K. Wijnrocx, and N. Gengler. 2022. Description of the genetic evaluation systems used in the Walloon Region of Belgium. Accessed Nov. 21, 2022. http://www.elinfo.be/indexEN.html.

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

VanRaden, P.M., J.B. Cole, and K.L. Parker Gaddis. 2018. Net merit as a measure of lifetime profit: 2018 revision. Accessed Aug. 01, 2022. https://www.aipl.arsusda.gov/reference/nmcalc-2018.htm.

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.

WHO and FAO (World Health Organization and Food and Agriculture Organization of the United Nations). 2011. Codex Alimentarius: Milk and Milk Products. 2nd ed. WHO FAO, Rome, Italy.

Wiggans, G.R., and J.A. Carrillo. 2022. Genomic selection in United States dairy cattle. Front. Genet. 13:1–7. https://doi.org/10.3389/fgene.2022.994466.

Wiggans, G.R., T.S. Sonstegard, P.M. VanRaden, L.K. Matukumalli, R.D. Schnabel, J.F. Taylor, F.S. Schenkel, and C.P. van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. J. Dairy Sci. 92:3431–3436. https://doi.org/10.3168/jds.2008-1758.

Wisconsin Alumni Research Foundation. 2015. Milk Urea-N yield (MUNY) as a nutritional and environmental management tool for the dairy industry. US Pat. No. 9,955,673 B1.

Wilmot, H., J. Bormann, H. Soyeurt, X. Hubin, G. Glorieux, P. Mayeres, C. Bertozzi, and N. Gengler. 2022. Development of a genomic tool for breed assignment by comparison of different classification models: Application to three local cattle breeds. J. Anim. Breed. Genet. 139:40–61. https://doi.org/10.1111/jbg.12643.

6

Chapter VI Genome-wide association study and functional annotation analyses for nitrogen use efficiency index and its composition in dairy cattle

Adapted from: Chen, Y., H. Atashi, C. Grelet, R.R. Mota, S. Vanderick, GplusE Consortium & N. Gengler. Genome-wide association study and functional annotation analyses for nitrogen use efficiency index and its composition in dairy cattle. J. Dairy Sci. Accepted.

Foreword

The new nitrogen efficiency index (NEI) was built in Chapter 4, and its ability to allow for genomic prediction was confirmed in Chapter 5. Furthermore, the genomic background of the NEI and its composition traits were analyzed through genome-wide association analyses in the present Chapter. This provided a better understanding of the genetic mechanisms of nitrogen use efficiency in dairy cows.

Abstract

The aims of this study were 1) to identify genomic regions associated with a nitrogen (N) efficiency index (NEI) and its composition traits; and 2) to analyze the functional annotation of identified genomic regions. The NEI included N intake (NINT1), milk true protein N (MTPN1), milk urea N yield (MUNY1) in primiparous cattle, N intake (NINT2+), milk true protein N (MTPN2+), and milk urea N yield (MUNY2+) in multiparous cattle (2 to 5 parities). The edited data included 1,043,171 records on 342,847 cows distributed in 1,931 herds. The pedigree consisted of 505,125 animals (17,797 males). Data of 565,049 SNPs were available for 6,998 animals included in the pedigree (5,251 females and 1,747 males). The SNP effects were estimated using a single-step genomic BLUP approach. The proportion of the total additive genetic variance explained by windows of 50 consecutive SNPs (with an average size of about 240 Kb) was calculated. The top-3 genomic regions explaining the largest rate of the total additive genetic variance of the NEI and its composition traits were selected for candidate genes identification and quantitative trait loci (QTLs) annotation. The selected genomic regions explained from 0.17% (MTPN2+) to 0.58% (NEI) of the total additive genetic variance. The largest explanatory genomic regions of NEI, NINT1, NINT2+, MTPN1, MTPN2+, MUNY1, MUNY2+ were Bos taurus autosome (BTA) 14 (1.52 – 2.09 Mb), 26 (9.24 – 9.66 Mb), 16 (75.41 – 75.51 Mb), 6 (8.73 – 88.92 Mb), 6 (8.73 – 88.92 Mb), 11 (103.26 – 103.41 Mb), 11 (103.26 – 103.41 Mb). Based on the literature, gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and protein-protein interaction (PPI), 16 key candidate genes were identified for NEI and its composition traits, which are mainly expressed in the milk cell, mammary, and liver tissues. The number of enriched QTLs related to NEI, NINT1, NINT2+, MTPN1, MTPN2+ were 41, 6, 4, 11, 36, 32, and 32, respectively, and most of them were related to the milk, health, and production classes. In conclusion, this study identified genomic regions associated with NEI and its composition traits, and identified key candidate genes describing the genetic mechanisms of NUE traits. Furthermore, the NEI not only reflects its composition traits, but also reflects the interaction among them.

Key words: gene, QTL, enrichment analysis

1. Introduction

High-efficiency dairy cattle are increasingly being pursued by milk producers (Brito et al., 2020). Cattle consume nitrogen (N) mostly in the form of feed crude protein, which is then degraded into different forms of N, such as amino acids, ammonia, and urea, for metabolism in the body (Aguirre-Villegas et al., 2017). The N emissions from the livestock sector accounted for one-third of current human-induced N emissions (Uwizeye et al., 2020). Dairy cows with high N use efficiency (NUE) not only improve the profitability of dairy farms, but also reduce environmental N pollution (Calsamiglia et al., 2010). The NUE is a complex trait involving multiple features, such as N intake (NINT), milk true protein N (MTPN), and milk urea nitrogen (MUN) (Chen et al., 2022). Milk urea concentration (MU) and MUN are the most commonly used NUE proxies in dairy cattle management and genetic breeding programs. The reason why MU (MUN) indirectly increases NUE is its strong correlation with urinary N (Kauffman and St-Pierre, 2001).

The traditional definition of NUE in dairy cows is milk N out divided by NINT, however, several shortcomings of this definition were shown in our latest article (Chen et al., 2022). Recently, we proposed a new N efficiency index (NEI) that considers both NUE and N pollution at the same time (Chen et al., 2022). The NEI is a combination of NINT, MTPN, and MUN yield (MUNY) being predicted by milk mid-infrared (MIR). The genetic correlations between NEI and production yield traits were positive, while the genetic correlation with the investigated functional traits were negative (Chen et al., 2022). However, the biological background of NEI is still missing. Although some studies have explained the biological background of MUNY (or MUN) (Strucken et al., 2012; Honerlagen et al., 2021; Ariyarathne et al., 2021), to our best knowledge, the biological backgrounds of NINT and MTPN have yet not been investigated. In addition, some studies performed genetic analyses on minor N compounds in milk (such as ammonia) and urinary urea (Pegolo et al., 2018; Honerlagen et al., 2021). However, these phenotypes are difficult to measure and are so far too challenging to be applied for dairy breeding purposes.

Multiple studies have shown that single-step genome-wide association study (ssGWAS) is an efficient method for studying the genomic background of complex traits (Li et al., 2019; Atashi et al., 2020; Brunes et al., 2021). Indeed, the ssGWAS algorithm can directly obtain the SNP variance through the genomic EBV (GEBV), allowing the estimation of the proportion of each SNP in the total additive genetic variance (Wang et al., 2012). However, the variance effect of a single SNP is often small, so it is a good way to express the proportion of genomic regions (SNP windows) of several consecutive SNPs in the total additive genetic variance (Fragomeni et al., 2014). The functional analysis of genes inside the identified genomic regions can better explain the genomic background of the research traits. For example, the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of genes located in genomic regions associated with a trait can reveal the biological process and pathway involved. Based on the genetic relationships between NEI and other traits at different strength levels (Chen et al., 2022), we speculate that the genetic region identified for NUE-related features may also regulate other traits. In addition, previous studies reported that the quantitative trait loci (QTLs) of MUNY were located on different chromosomes, showing the polygenic profile of this trait (Bouwman et al., 2010; Strucken et al., 2012). Identified genomic regions can be compared with the QTL (genomic regions) previously reported and checked for the potential impact of genetic selection of NUE on other traits at the QTL level.

The objectives of this study were to investigate the genomic background of the NEI and verify whether the NEI can reflect the combined three NUE-related features. In this regard, ssGWAS was used to identify genomic regions associated with NEI and its composition traits; then, functional annotation analyses were performed on the genomic regions identified for the corresponding traits.

2. Materials and methods

The study framework is shown in Figure 6-1.

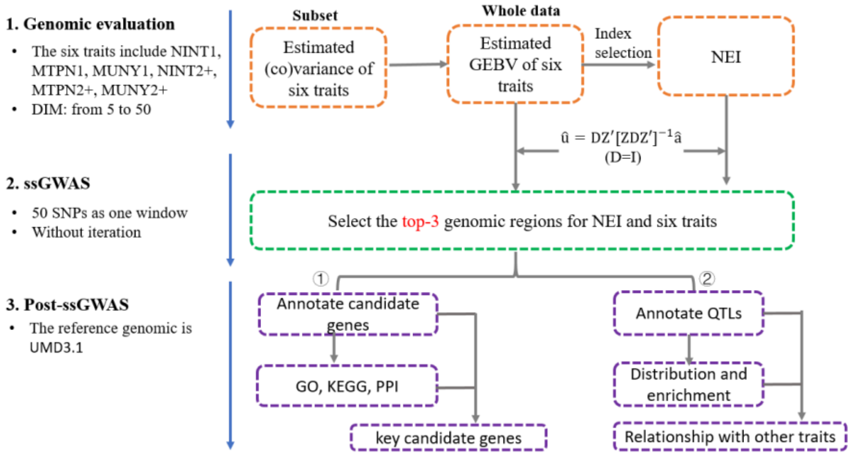


Figure 6-1. Workflow for the N efficiency index (NEI), the N intake in primiparous cows (NINT1), N intake in multiparous cows (NINT2+), milk true protein N in primiparous cows (MTPN1), milk true protein N in multiparous cows (MTPN2+), milk urea N yield in primiparous cows (MUNY1), and milk urea N yield in multiparous cows (MUNY2+)

2.1. Data

Phenotypic Data: the data used in this study was the same as that used by Chen et al. (2022). Briefly, we used 1,043,171 test-day records, collected from 2001 to 2019 on 342,847 cows distributed in 1931 herds. The range of days in milk (DIM) for all used records was restricted to between 5 and 50 because of the predicament of the NINT model. The NINT of each cow was predicted by the equations based on the models developed by Grelet et al. (2020). The following formulas were used to compute MTPN = [(Milk yield × Protein percentage / 6.38) – MUNY] and MUNY = [(Milk urea concentration / 2.14) × Milk yield] (WHO and FAO, 2011) with protein percentage and milk urea concentration also predicted by FT-MIR analysis of milk. In addition, we divided the 3 NUE-related features (NINT, MTPN, MUNY) into NINT1, MTPN1, MUNY1, NINT2+, MTPN2+, and MUNY2+ traits according to primiparous and multiparous classes (2 to 5 parities). The pedigree related to the dataset comprised 505,125 animals (17,797 males). Grelet et al. (2020) used support vector machine regression to build a NINT prediction model based on milk MIR spectra, milk yield, and parity. Milk MIR spectra of 53,660 cows from 776 herds were available, and the number of NINT records was 143,595.

Genotypic data: Genotypic data was available for 6,998 animals (1,747 males and 5,251 females). Individuals were genotyped using the BovineSNP50 Beadchip v1 to v3 (Illumina, San Diego, CA, USA). Single nucleotide polymorphisms (SNP) common between all the three chips were kept. Non-mapped SNPs, SNP located on sex chromosomes, and triallelic SNPs were excluded. A minimum GenCall Score of 0.15 and a minimum GenTrain Score of 0.55 were used to keep SNP (Wilmot et al., 2022). The genotypes were imputed to HD by using FImpute V2.2 software (Sargolzaei et al., 2014). One of the common editing steps for marker data (e.g. SNP) is to check for Mendelian conflicts (Wiggans et al., 2009). A Mendelian conflict occurs when the genotype and pedigree data of two related animals are in disagreement. This may result from an error in the recorded pedigree, from genotyping errors, from mixing up DNA samples and in very rare cases from mutations (Calus et al., 2011). In this study, SNPs with Mendelian conflicts and those with minor allele frequency less than 5% were excluded. The difference between observed heterozygosity and that expected under Hardy-Weinberg equilibrium was estimated, and SNPs difference greater than 0.15 were excluded (Wiggans et al., 2009). In total, 565,049 SNPs located on 29 Bos taurus autosomes (BTA) were used in the genomic analyses.

2.2. (Co)variance Components Estimation

In total, 143,595 test-day records on 53,660 cows for six traits extracted from the whole dataset were used to estimate (co)variance components (Chen et al., 2022). The pedigree of (co)variance components data included 133,943 animals (7,879 males), and 4,563 (1,292 males) of them with genotype. A six-trait (six traits; 3 traits and 2 parity classes) repeatability model was used to estimate the variance components. The information of the model can be found in Chen et al. (2022). In brief, herd-year-season of calving, DIM, and calving age (nested within parities) were used as fixed effects in this model, whereas non-genetic cow, non-genetic cow × parity (only for multiparous traits), additive animal genetic, and residual were used as random effects. However, when calculating the relationship between animals, we used the **H** matrix, which combined pedigree(**A**) and genomic (**G**) relationships. The inverse of H as defined by Aguilar et al. (2010):

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

where **A** was the numerator relationship matrix based on the pedigree; $\mathbf{A}_{22}$ was the numerator relationship matrix based on the pedigree for genotyped animals; **G** was the genomic relationship matrix which was obtained using the function described by VanRaden (2008). In addition, it should be noted that the inverse of all matrices considers the coefficient of inbreeding between individuals (Lourenco et al., 2020).

Computations were performed using the BLUPF90 family of programs (Misztal et al., 2018). The (co)variance components and parameters for NINT, MTPN, and MUNY were estimated by Gibbs sampling, which was same as the Chen et al. (2022). The formulas used to calculate the parameters heritability (h^2), repeatability, genetic and phenotypic correlation) were the same as that of Chen et al. (2021).

2.3. Estimated Breeding Values and Nitrogen Efficiency Index

Using the above estimated genetic variance component and related model (equation 1), GEBV were estimated for 6 traits, according to the precondition conjugate algorithm implemented in the BLUPF90 (version 1.71) program. The whole dataset was used for this purpose. The calculation method of NEI was the same as Chen et al. (2022). In short, The NEI was obtained by combining the GEBVs of the six traits using the selection index theory. The relative weights of the six considered traits were calculated by selection responses, which assumed that the selection responses for NINT, MTPN, and MUNY were 0, 1, and -1, respectively.

2.4. Genome-wide Association Analyses

The SNP effects for the NEI and six traits were estimated using the POSTGSF90 software (version 1.73; Aguilar et al., 2014). The formula used for estimating SNP effects was as follows Wang et al. (2012):

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

where $\hat{\mathbf{a}}$ is the SNP effect, **D** = **I**, which means the weight for all SNPs is 1; $\mathbf{Z}_{\mathbf{g}}$ was an incidence matrix of genotyped for each SNP; $\hat{\mathbf{u}}$ was a vector of GEBV for each trait of genotyped animals**.** The variance of i SNP is $\hat{d}_i = a_i^2 2p_i(1-p_i)$, where a_i^2 is the square of *i*th SNP solution, and p_i is the frequency of allele B at SNP *i*. The results were presented by the proportion of variance explained by each window of 50 adjacent SNPs with an average size of 240 Kb. We used 1 SNP as the moving step of the window, which ensured that we do not miss genomic regions potentially associated with the trait due to the combination of SNPs. The formula for the additive genetic variance of each window was as follows:

$$\frac{\mathrm{var}(a_i)}{\sigma_a^2} \times 100\ (\%) = \frac{\mathrm{var}(\sum_{j=1}^{50} Z_j\hat{d}_j)}{\sigma_a^2} \times 100\ (\%)$$

where a_i is the genetic variance of the *i*th genomic region (each window combines 50 SNP), σ_a^2 is the total genetic variance, Z_j is the vector of the SNP content of the *j*th SNP for all individuals, and $\hat{d}_j$ is the variance of the *j*th SNP.

Linkage disequilibrium (squared correlation coefficient, r^2) was calculated for SNPs within a window that explained more than 0.5% of additive genetic variance.

2.5. Functional Annotation Analyses

Following Soares et al. (2021), the top genomic regions were selected to investigate candidate genes and their annotation. However, due to the large number of traits considered and the small proportion of variance explained by genomic regions in the current study, only the top-3 genomic regions were selected. Then, candidate genes and QTL annotations were performed through the GALLO R package (Fonseca et al., 2020).

Identification of protein-encoding genes located in these selected genomic regions, using the Bos taurus UMD3.1.94 assembly as the reference map (http://ftp.ensembl.org/pub/release-94/gtf/bos_taurus/, accessed on October 19, 2021). The gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were carried out on the candidate gene sets obtained for NEI and included six traits through the clusterProfiler R package (Wu et al., 2021). Furthermore, the protein-protein interaction (PPI) analysis was performed on the candidate genes obtained from the analyzed traits through STRING (Szklarczyk et al., 2021) to reveal the relationship between the identified candidate genes. The PPI relationship was based on text mining, experiments, database, co-expression, neighborhood, gene fusion, and co-occurrence, and the minimum required interaction score was set to 0.4 (Zhou et al., 2019). Then, the Cytohubba in Cytoscape (version 3.8.2) was used to find the hub genes. Based on the literature, GO, KEGG, and PPI, important candidate genes were selected, hereafter referred to as key candidate genes. Moreover, we checked the expression levels of the candidate genes (or key genes) over 100 tissues/cell types in cattle through the cGTEx database (https://cgtex.roslin.ed.ac.uk/; Liu et al., 2022).

The top-3 genomic regions of the studied traits were annotated with Cattle QTLdb (UMD3.1, https://www.animalgenome.org/cgi-bin/QTLdb/BT/inde, accessed on October 19, 2021; Hu et al., 2019). At present, Cattle QTLdb has 158,041 QTLs, which were divided into six 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 the enrichment analysis (Fonseca et al., 2020). In all enrichment analyses (GO, KEGG, QTL), the Benjamini-Hochberg method was used for multiple testing corrections.

3. Results and discussion

3.1. Genetic Parameter

The genetic parameters estimated for the studied NUE-related traits were described in Table 6-1, which were similar to our previous results without using genotypes (Chen et al., 2022). The h^2 and repeatability for the six NUE-related traits ranged from 0.10 to 0.14, and 0.38 to 0.64, respectively. Compared with our previous study (Chen et al., 2022), the change in the variance component ranged from −3.00% to 0.88%; the absolute value of h^2 changed from 4.64% (MUNY1) to 8.81% (MUNY2+), and of repeatability, from 0.36% (NINT2+) to 3.32% (NINT1). Therefore, it can be concluded that including genotypic data in the variance components analysis showed a minor impact and the studied NUE-related traits have low h^2 and medium repeatability. Although we do expect the genotypes to capture a bit more variance that increases the h^2, especially for complex traits, the genotype population in this study is small (n < 5,000) (de Los Campos et al., 2018), and substantial changes were not observed.

Table 6-1. Heritability (h^2), repeatability, additive genetic variance (σ^2_a), across-parity permanent environment (non-genetic cow) variance (σ^2_c), (only for second and later lactations) within parity permanent environment (non-genetic cow x parity) variance (σ^2_p), and residual variance (σ^2_e) of the proxies for predicted N intake (NINT, expressed as 100 g/day), milk true protein N (MTPN, expressed as 100 g/day), and yield of milk urea N (MUNY, g/day) in primiparous (n = 44,321) and multiparous (n = 99,274) Holstein cows

Trait[1]	h^2	repeatability	σ^2_a	σ^2_c	σ^2_p	σ^2_e
NINT1	0.13 ± 0.01	0.38 ± 0.01	0.03 ± 0.00[2]	0.06 ± 0.00	NA[3]	0.14 ± 0.00
MTPN1	0.12 ± 0.00	0.59 ± 0.00	0.01 ± 0.00	0.02 ± 0.00	NA	0.02 ± 0.00
MUNY1	0.14 ± 0.01	0.40 ± 0.01	0.10 ± 0.01	0.20 ± 0.01	NA	0.45 ± 0.01
NINT2+	0.12 ± 0.01	0.45 ± 0.00	0.04 ± 0.00	0.03 ± 0.00	0.08 ± 0.00	0.17 ± 0.00
MTPN2+	0.10 ± 0.01	0.64 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.03 ± 0.00	0.03 ± 0.00
MUNY2+	0.10 ± 0.01	0.43 ± 0.00	0.16 ± 0.01	0.13 ± 0.01	0.40 ± 0.01	0.93 ± 0.01

[1] Trait: NINT1 - N intake in primiparous cows; MTPN1 - milk true protein N in primiparous cows; MUNY1 - Milk urea nitrogen yield in primiparous cows; NINT2 - N intake in multiparous cows; MTPN2 - milk true protein N in multiparous cows; MUNY2 - Milk urea nitrogen yield in multiparous cows

[2] Standard error is less than 0.005

[3] NA: not applicable

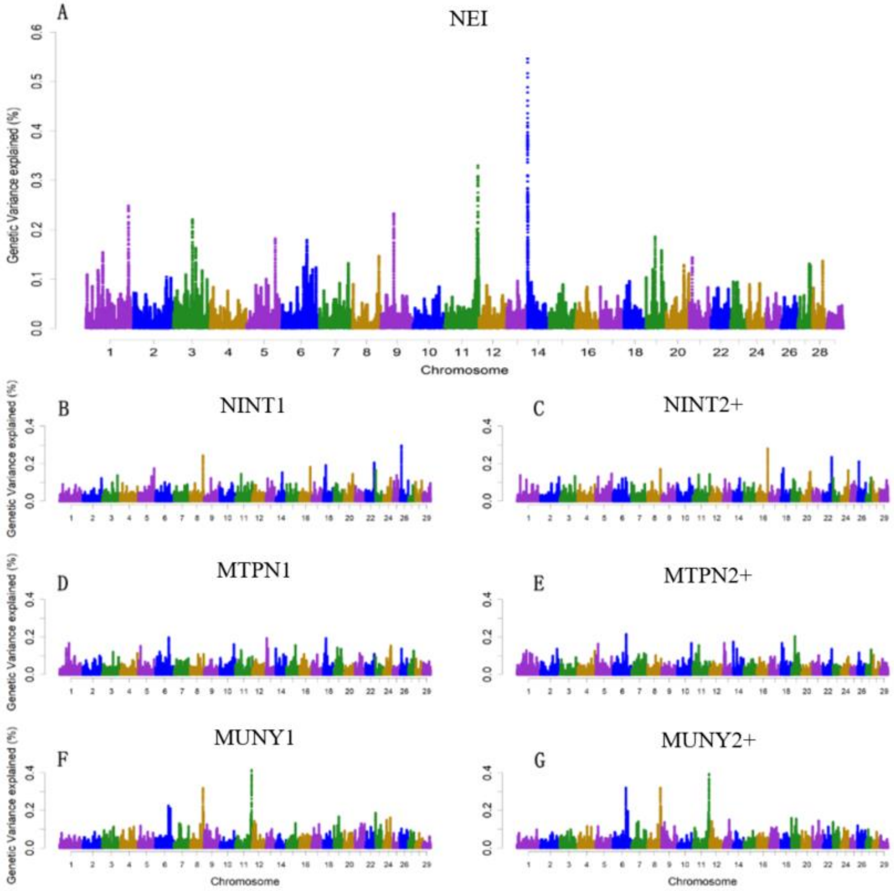


Figure 6-2. Additive genetic variance was explained by windows of 50 adjacent SNP across chromosomes for the N efficiency index (NEI, A), the N intake in primiparous cows (NINT1, B), N intake in multiparous cows (NINT2+, C), milk true protein N in primiparous cows (MTPN1, D), milk true protein N in multiparous cows (MTPN2+, E), milk urea N yield in primiparous cows (MUNY1, F), and milk urea N yield in multiparous cows (MUNY2+, G)

3.2. Genome-wide Association Analyses

The Manhattan plot for the studied traits were shown in Figure 6-2. The top-3 genomic regions selected for the studied traits were shown in Table 6-2. The genomic regions identified for NEI, NINT, MTPN, MUNY were in Bos Taurus autosome (BTA) 11, 14; 8, 16, 22, 25, 26; 6, 14, 13, 18, 19; 6, 8, 11. Although the genomic regions identified by the same feature in the two parity classes were not the same, the genomic regions identified were similar, especially for MUNY. Soares et al. (2021) found a similar situation when they studied clinical ketosis. The genomic regions identified for NEI were also associated with MTPN and MUNY (except the third region), which confirmed that NEI is affected by MTPN and MUNY. The identified genomic regions explained from 0.17% (MTPN2+) to 0.58% (NEI) of the total additive genetic variance.

The top-3 genomic region combined explained 1.18%, 0.75%, 0.73%, 0.58%, 0.59%, 0.96%, and 1.03% of the total additive genetic variance for NEI, NINT1, NINT2+, MTPN1, MTPN2+, MUNY1, and MUNY2+, respectively. Results showed that most windows explained less than 0.50% of the total additive genetic variance of the traits, and these low contributing regions spread across the genome for all traits analyzed. It indicates that NEI and NUE-related traits are moderate to highly polygenic, in which many regions across the genome contribute to the genetic variation of the traits. Similar results were reported for MUNY by Strucken et al (2012). It should be noticed that windows explained more than 0.50% of the total additive genetic variance identified only for NEI, which means that NEI may reflect interactive effects between MTPN and MUNY. The linkage disequilibrium for genomic regions greater than 0.50% of the NEI was shown in Figure 6-3, with 5 blocks shown inside. We found that 38.89% (7/18) of the genes in the nitrogen metabolism pathway (KEGG:00910) were located in BTA14 (27.63 to 79.73 Mb), which included *CA1*, *CA2*, *CA3*, *CA8*, *CA13*, *LOC784254*, and *LOC100847874*. From this point, only the genomic regions that explained the highest genetic variance of the studied traits were discussed.

Table 6-2. Annotated genes within the top-3 genomic regions explaining the biggest proportion of genetic variance for N efficiency index (NEI)

Trait	BTA[1]	Position(bp)[2]	Var[3]	Genes[4]
NEI	14	1517553: 2089613	0.58	*ZNF7, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT, PPP1R16A, FOXH1, KIFC2, CYHR1, TONSL, VPS28, SLC39A4, CPSF1, ADCK5, SLC52A2, FBXL6, TMEM249, SCRT1, DGAT1, HSF1, BOP1, SCX, MROH1, HGH1, WDR97, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, SPATC1, GRINA, PARP10, PLEC*
	11	103264921: 103409247	0.33	*PAEP, GLT6D1, LCN9, KCNT1*
	14	2673388: 2978629	0.27	*GML, LY6K, LY6D, LYNX1, LYPD2, SLURP1, THEM6, PSCA, ARC, ADGRB1*
NINT1	26	9242669:9655433	0.30	*PAPSS2, ATAD1, PTEN*
	8	103696313:103829659	0.24	*SNX30, SLC46A2*
	22	55490915:55638639	0.20	*SLC6A11*
NINT2+	16	75405390:75509546	0.28	*IRF6, C16H1orf74, TRAF3IP3, HSD11B1*
	22	55490915:55638639	0.23	*SLC6A11*
	26	9242669:9655433	0.21	*PAPSS2, ATAD1, PTEN*
MTPN1	6	88732184:88919352	0.20	*GC*
	13	10175391:10315354	0.19	*KIF16B*
	18	15797080:15884324	0.19	*ITFG1*
MTPN2+	6	88732184:88919352	0.22	*GC*
	19	22594096:22657020	0.20	*NXN, MRM3, GLOD4, DBIL5;*
	14	1517553:2089613	0.17	*ZNF7, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT, PPP1R16A, FOXH1, KIFC2, CYHR1, TONSL, VPS28, SLC39A4, CPSF1, ADCK5, SLC52A2, FBXL6, TMEM249, SCRT1, DGAT1, HSF1, BOP1, SCX, MROH1, HGH1, WDR97, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, SPATC1, GRINA, PARP10, PLEC*
MUNY1	11	103264921:103409247	0.41	*PAEP, GLT6D1, LCN9, KCNT1*
	8	103694244:103828116	0.32	*INIP, SNX30, SLC46A2*
	6	87136725:87296185	0.22	*CSN1S1, CSN2, HSTN, STATH, CSN1S2*
MUNY2+	11	103264921:103409247	0.39	*PAEP, GLT6D1, LCN9, KCNT1*
	6	87145250:87311202	0.32	*CSN1S1, CSN2, HSTN, STATH, CSN1S2*
	8	103696313:103829659	0.32	*SNX30, SLC46A2*

[1]: BTA = Bos taurus autosomes

[2]: Starting and ending coordinates of the genomic region

[3]: Var = percentage of genetic variance explained by the SNPs within the genomic region

[4]: Genes: EBSEMBL symbol of annotated genes using the Bos Taurus UMD3.1.94 assembly (http://ftp.ensembl.org/pub/release-94/gtf/bos_taurus/)

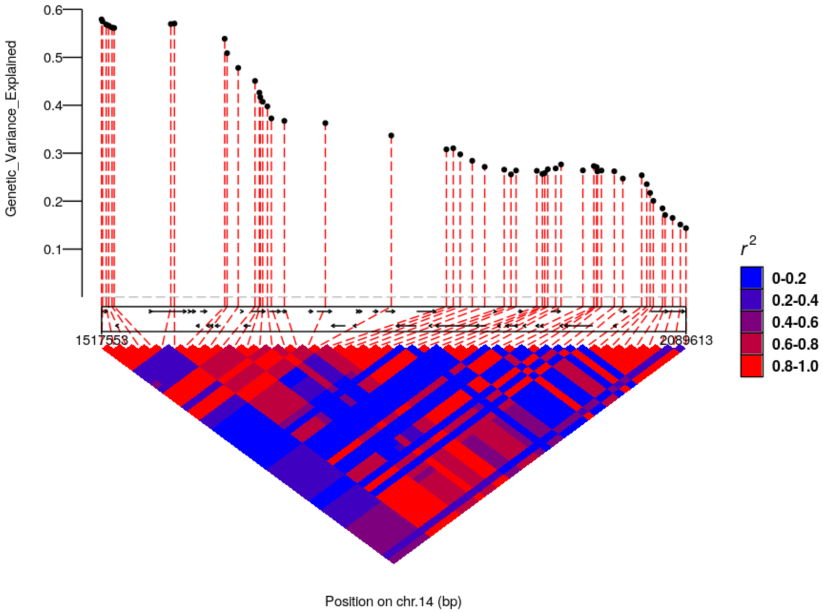


Figure 6-3. Linkage disequilibrium between 50 SNPs inside the genomic region on BTA14 position 1.52 to 2.09 Mb associated with nitrogen efficiency index

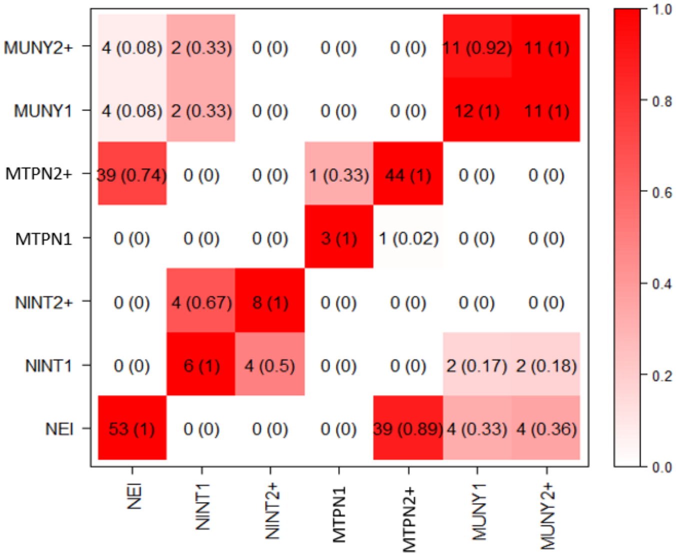
The genomic region between 1.52 to 2.09 Mb on BTA14 explained 0.58% of the total additive genetic variance of NEI. This region has been reported to be associated with milk yield by multiple studies (Nayeri et al., 2016; Atashi et al., 2020; Bakhshalizadeh et al., 2021). At the same time, this region (BTA14: 1.52 to 2.09 Mb) was also associated with protein and fat yields (Veerkamp et al., 2016; Cai et al., 2019). This is in line with our prediction of improving NEI through milk true protein N. The genomic regions between 9.24 to 9.66 Mb on BTA26 and between 75.41 to 75.51 Mb on BTA16 explained the highest part of the total additive genetic variance of NINT1 and NINT2+, respectively. The SNP inside the genomic region found on BTA26 position 9.24 to 9.66 Mb were associated with milk yield, milk C14 index, and milk myristoleic acid content (Minozzi et al., 2013; Gebreyesus et al., 2019). The genomic region identified on BTA16 position 75.41 to 75.51 Mb was associated with the residual feed intake and feed efficiency in cattle (Brunes et al., 2021). The genomic region between 88.73 to 88.92 Mb on BTA 6 accounted for the highest total additive genetic variance of MTPN1 and MTPN2+. This genomic region was associated with protein yield and its composition (Olsen et al., 2016; Zhou et al., 2019). The genomic region between 103.26 to 103.41 Mb on BTA11 was associated with MUNY1 and MUNY2+. This region has been reported to be associated with MU in Brown Swiss cattle (Pegolo et al., 2018), with MUN in Australian and New Zealand dairy cattle (van den Berg et al., 2022a). Ariyarathne et al. (2021), using the 50K SNP chip, found that the position 100 to 101 Mb on BTA11 was associated with MU in mixed-breed of cattle (Holstein Friesian (F), and Jersey(F), F x J crossbred). Previous studies showed that the genomic region on BTA6 position 87 to 88 Mb was associated with MU (Pegolo et al., 2018; Ariyarathne et al., 2021). This region was among the identified top-3 genomic regions associated with MUNY. Previous studies reported the genomic region between 6.12 to 7.15 was associated with MU yield (Honerlagen et al., 2021; Strucken et al., 2012). The differences observed in the ongoing study can be explained by the smaller number of cows used in their studies (i.e. 371 and 152, respectively).

Briefly, the genomic regions identified for the studied traits were located on multiple BTA and explained a small fraction of the total additive genetic variance, suggesting these traits are complex quantitative traits controlled by multiple genes. Increasing the SNP density in these genomic regions of NEI and six considered traits (especially NINT and MUNY) when making the SNP chip may improve the reliability of genetically selected for NUE. The NEI may better reflect NUE because it has a prominent peak at BTA14, which is closer to genes related to nitrogen metabolism pathways.

3.3. Gene Annotation Analyses

The results of gene annotation analyses are described in Table 6-2 and Figure 6-4. There were no common candidate genes between NEI and NINT (Figure 6-4), which is consistent with our hypothesis (keeping NINT unchanged). Surprisingly, there was no common candidate gene between NEI and MTPN1, which could be that NEI only increases the NUE of dairy cows through MTPN2+ in the first five parities. The percent of common candidate genes for NEI and MTPN2+, MUNY ranged from 33% to 89%.

Among the annotated candidate genes, *DGAT1, GRINA, CYHR1, FOXH1, TONSL, PPP1R16A, ARHGAP39, MAF1,* OPLAH, *MROH1, ZNF7, SLURP1, MAFA, KIFC2, GML, PSCA, THEM6, LYNX1,* and *ARC* have been reported to be associated with to 305-d milk yield (Nayeri et al, 2016; Atashi, et al., 2020). The *DGAT1* was also reported as a gene associated with milk yield, fat, and protein percentages (Bakhshalizadeh et al. (2021)). The *CSN1S1, CSN1S2, CSN2,* and *PAEP* have been reported to be associated with milk protein composition (Sanchez et al., 2017; Pegolo et al., 2018; Zhou et al., 2019). The *BOP1* was associated with protein yield (Cai et al., 2019). Brune et al. (2021) reported that the *MAF1* was associated with low animal feed intake.



	NEI	NINT1	NINT2+	MTPN1	MTPN2+	MUNY1	MUNY2+
MUNY2+	4 (0.08)	2 (0.33)	0 (0)	0 (0)	0 (0)	11 (0.92)	11 (1)
MUNY1	4 (0.08)	2 (0.33)	0 (0)	0 (0)	0 (0)	12 (1)	11 (1)
MTPN2+	39 (0.74)	0 (0)	0 (0)	1 (0.33)	44 (1)	0 (0)	0 (0)
MTPN1	0 (0)	0 (0)	0 (0)	3 (1)	1 (0.02)	0 (0)	0 (0)
NINT2+	0 (0)	4 (0.67)	8 (1)	0 (0)	0 (0)	0 (0)	0 (0)
NINT1	0 (0)	6 (1)	4 (0.5)	0 (0)	0 (0)	2 (0.17)	2 (0.18)
NEI	53 (1)	0 (0)	0 (0)	0 (0)	39 (0.89)	4 (0.33)	4 (0.36)

Figure 6-4. The number (percentage) of share genes among N efficiency index (NEI), N intake in primiparous cows (NINT1), N intake in multiparous cows (NINT2+), milk true protein N in primiparous cows (MTPN1), milk true protein N in multiparous cows (MTPN2+), milk urea N yield in primiparous cows (MUNY1), and milk urea N yield in multiparous cows (MUNY2+). PS: the upper left triangle is the number of candidate genes for NEI as the denominator and the lower right triangle is the number of candidate genes for traits as the denominator

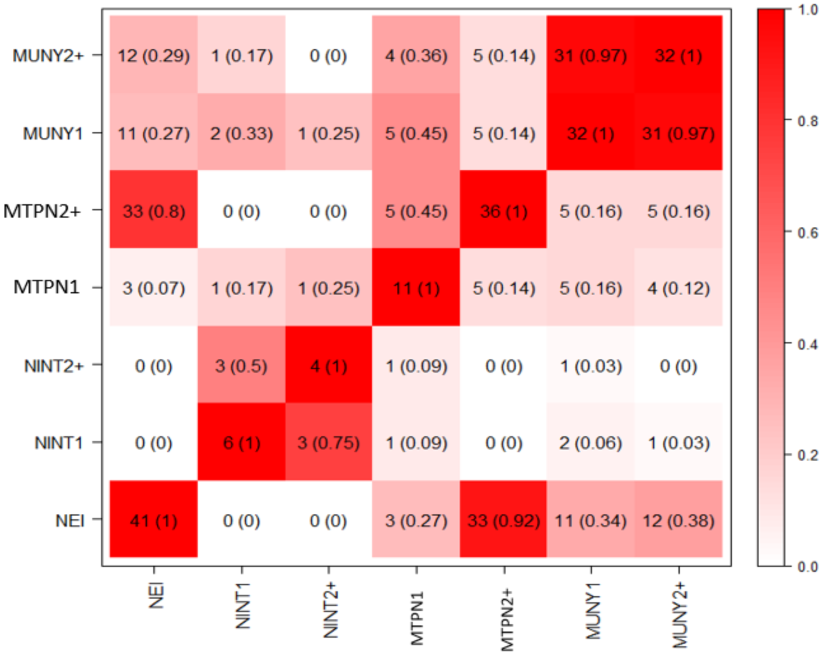
The results of the GO enrichment analysis were presented in Supplement file 1. The candidate genes identified for NINT1, NINT2+, MTPN1, MUNY1, and MUNY2+ enriched 33, 5, 3, 23, 28, and 28 GO terms, however, the candidate genes identified for NEI and MTPN2+ enriched no GO term. The 33 GO terms enriched by the NINT only involve *ATAD1*, which was also identified as a candidate gene for dry matter intake (Serão et al., 2013). The MTPN1 got the 3 GO terms by *GC*, which encodes the Vitamin D Binding protein. The *GC* gene was identified as being associated with milk production, mastitis, and postpartum blood calcium concentration (Olsen et al., 2016; Cavani et al., 2022). The 28 GO terms enriched of the MUNY through *CSN1S1*, *CSN1S2*, *CSN2*, *HSTN*, and *STATH*. The first three genes (*CSN1S1*, *CSN1S2*, *CSN2*) belong to CSN@ (casein cluster) family genes, which were identified as related to α_{s1}-casein, β-casein, and k-casein (Zhou et al., 2019). The *HSTN* was identified as related to k-casein (Zhou et al., 2019), which also affected the β-casein and α_{s2}-casein (Elsik et al., 2009). The *STATH* affected the β-casein (Rijnkels et al., 2003).

Only MUNY1 and MUNY2 enriched 2 pathways: salivary secretion (bta04970) and prolactin signaling pathway (bta04917) (Supplemental file 2). These pathways were enriched by *CSN2*, *HSTN*, and *STATH*, and the 3 genes were explained the same as in the GO analysis part. The salivary secretion impacts the feed intake in cattle (Taussat et al., 2020). The salivary secretion also is associated with methane emissions in non-lactating dairy cows (Pinares-Patiño et al., 2007). The prolactin signal regulates the milk production and composition of dairy cows (Raven et al., 2014).

The internal gene interaction relationships of the candidate gene sets for NEI and its composition traits were shown in supplementary file 3 (from Figure 6-S1 to 6-S5). The PPIs for NEI, NINT1, NINT2+, MTPN1, MTPN2+, MUNY1, and MUNY 2+ were composed of 58 nodes and 188 edges, 11 nodes and 8 edges, 13 nodes and 9 edges, 8 nodes and 15 edges, 49 nodes and 156 edges, 17 nodes and 32 edges, 16 nodes and 41 edges, respectively. And the PPI enrichment *P* of NEI, MTPN, and MUNY were less than $4 \times 1.0E^{-4}$. The NEI had a similar PPI to MTPN2+, which was also similar to the PPI of dairy components (Bakhshalizadeh et al., 2021). This is possibly caused by the candidate genes annotated in BTA14 1.52 to 2.09 Mb. Both the PPIs of MUNY1 and MUNY2+ showed the protein network in the STRING database (CL: 24892), which was composed of *CSN1S1*, *CSN1S2*, *CSN2*, *CSN3*, and *PAEP*. The *PAEP* encodes β-lactoglobulin, and the first four genes encode casein. The above 5 genes had been subject to long-term selection and changed their SNP frequencies in the cattle (Kolenda and Sitkowska, 2021). The TOP-1 of hub genes for NEI, NINT2+, MTPN1, MTPN2+, MUNY1, and MUNY2+ were *MROH1*, *PTEN*, *AHSG*, *MROH1*, *CSN1S1*, and *CSN2*. However, the PPI of NINT1 didn’t have hub genes. The *MROH1*, *CSN1S1*, and *CSN2* affected the milk protein composition (Sanchez et al., 2017). The activation of *PTEN* was not conducive to the lactation of dairy cows and reduced the production of β-casein (Wang et al., 2014).

In short, NEI and MTPN were affected by genes that control milk yield and milk production components, such as *DGAT1*, *PPP1R16A*, *CYHR1*, *CPSF1*, *MROH1*, *GC*, and *AHSG*. The NINT was affected by *ATAD1* and *PTEN* genes. The MUNY was affected by genes that control the production of casein (*CSN1S1*, *CSN1S2*, *CSN2*, *CSN3*, *HSTN*, *STATH*, and *PAEP*) and was related to the salivary secretion (bta04970). The expressions of the 16 key candidate genes involved in the cattle tissues were given in Supplemental file 4, and these genes were mainly expressed in milk cell, mammary, and liver tissues. The results of these genes and their high-expression tissue locations can be used for future studies on the genetic mechanisms of NUE.

3.4. QTL Annotation for Select Genomic Regions



	NEI	NINT1	NINT2+	MTPN1	MTPN2+	MUNY1	MUNY2+
MUNY2+	12 (0.29)	1 (0.17)	0 (0)	4 (0.36)	5 (0.14)	31 (0.97)	32 (1)
MUNY1	11 (0.27)	2 (0.33)	1 (0.25)	5 (0.45)	5 (0.14)	32 (1)	31 (0.97)
MTPN2+	33 (0.8)	0 (0)	0 (0)	5 (0.45)	36 (1)	5 (0.16)	5 (0.16)
MTPN1	3 (0.07)	1 (0.17)	1 (0.25)	11 (1)	5 (0.14)	5 (0.16)	4 (0.12)
NINT2+	0 (0)	3 (0.5)	4 (1)	1 (0.09)	0 (0)	1 (0.03)	0 (0)
NINT1	0 (0)	6 (1)	3 (0.75)	1 (0.09)	0 (0)	2 (0.06)	1 (0.03)
NEI	41 (1)	0 (0)	0 (0)	3 (0.27)	33 (0.92)	11 (0.34)	12 (0.38)

Figure 6-5. The number (percent) of enrichment QTLs among N efficiency index (NEI), N intake in primiparous cows (NINT1), N intake in multiparous cows (NINT2+), milk true protein N in primiparous cows (MTPN1), milk true protein N in multiparous cows (MTPN2+), milk urea N yield in primiparous cows (MUNY1), and milk urea N yield in multiparous cows (MUNY2+). PS: the upper left triangle is the number of candidate genes for NEI as the denominator and the lower right triangle is the number of candidate genes for traits as the denominator

The number of previously reported QTLs located in the identified genomic regions for the studied traits were described in Supplemental file 5 (Figure 6-S6). The proportion of selected QTLs in the QTL classification and the top-10 QTL of the milk class were in Supplemental file 5 (from Figure 6-S7 to 6-S13). But different traits had a different number of annotation studies, which made bias for annotation results (Fonseca et al., 2020). Therefore, based on the annotated QTLs, the enrichment analyses on NEI and the six studied traits. The number and class of significantly related QTLs were shown in Figure 6-5 and Figure 6-6. The percentage of common significant QTLs of NEI and MTPN1, MTPN2+, MUNY1, MUNY2+ were 3%, 33%, 11%, and 12%, respectively, which is also similar to our results of constructing NEI and annotating genes. The significant QTLs of NEI and MTPN were only distributed in milk, health, and production classes, except that the MTPN1 had one QTL in the exterior class. If the NEI was used in genetic selection, it means that NEI will affect these three trait classes, which is consistent with the genetic correlations between NEI and the traits (Chen et al., 2022). The significant QTLs annotated by NINT were distributed in the milk and production class. The significant QTLs annotated by MUNY were distributed in 4 classes of QTL, which may indicate the complexity of MUNY.

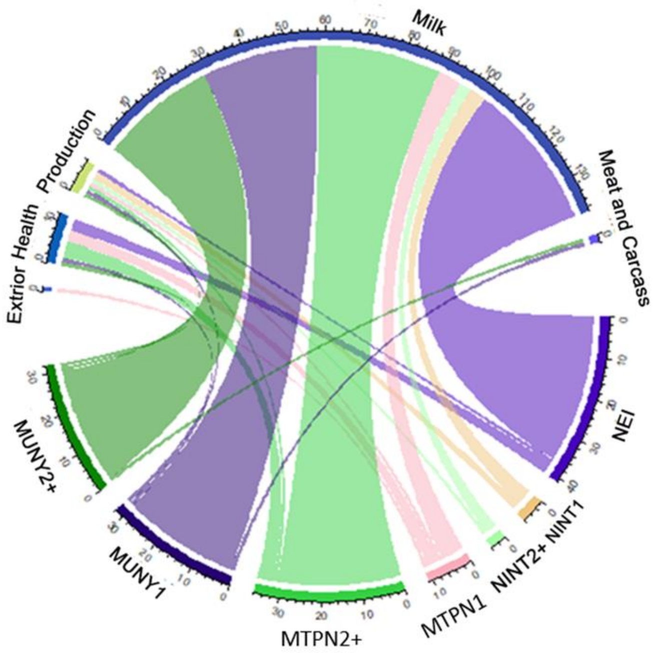


Figure 6-6. The number of enrichment QTLs class among N efficiency index (NEI), N intake in primiparous cows (NINT1), N intake in multiparous cows (NINT2+), milk true protein N in primiparous cows (MTPN1), milk true protein N in multiparous cows (MTPN2+), milk urea N yield in primiparous cows (MUNY1), and milk urea N yield in multiparous cows (MUNY2+).PS: QTLs classes were defined by cattle QTL database (https://www.animalgenome.org/cgi-bin/QTLdb/BT/ontrait?class_ID=1)

The top-10 QTLs after QTL enrichment analysis of the studies traits were shown in Supplemental file 6 (from Figure 6-S14 to 6-S20). The most significant related QTLs for NEI and MTPN2+ were those reported for milk yield which is located on BTA14. This genomic region on BTA14 position 1.52 to 2.09 Mb has been reported to be associated with milk production (Nayeri et al., 2016; Atashi et al., 2020; Bakhshalizadeh et al., 2021). We also found the lifetime profit index was related to the NEI, which explained the lifetime profit index related to the BTA14 1.6 to 1.8 Mb genomic regions (Nayeri et al., 2017). The most significant QTL for MTPN1 was also associated with clinical mastitis and located on BTA6 position 88.73 to 88.92 Mb

(Olsen et al., 2016, Freebern et al., 2020). The most significant QTL for NINT1 and NINT2+ has also been associated with the milk C14 index. The C14 index was identified as related to the genomic region (from 14.9 to 24.9 Mb) on BTA26 (Li et al., 2015), which was similar to the NINT (BTA26 position 9.24 to 9.66 Mb). The most significant correlation QTL for MUNY1 and MUNY2+ was milk beta-lactoglobulin percentage, which may be because BTA11 103.26 to 103.41 Mb was simultaneously reported to be related (Sanchez et al., 2017). The BTA11 position 103.26 to 103.41 Mb also was associated with MU and milk protein components (Ariyarathne et al., 2021; Pegolo et al., 2018). The BTA17 identified by MUNY had also been reported to be related to MU yield and Milk N (Honerlagen et al., 2021). Other studies had reported that the QTLs of MUNY were located on BTA3, BTA6, and BTA21 (Bouwman et al., 2010; Strucken et al., 2012).

In summary, the related QTLs of NEI were mainly reported for milk yield, fat, and protein composition. The related QTL of NINT was reported for the milk C14 index. The related QTLs of MTPN1 and MTPN2+ were related to clinical mastitis and milk yield, respectively. The related QTLs of MUNY were reported for protein composition.

The traits in this study were all related to traits predicted by milk MIR. The effect of using traits predicted from milk MIR has disadvantages and advantages for the impact on GWAS results. The disadvantage is that like other predictive traits, the accuracy of the prediction equation can have a large impact. If the accuracy of the prediction equation is very low, it may result in GWAS not working on the traits we want. The advantage of milk MIR prediction traits is that they can be predicted cheaply and at a large scale. Large-scale data are useful to overcome the problem of inaccurate GWAS results from small samples. Recently, van den Berg et al. (2022b) showed that using the blood urea nitrogen predicted by MIR increased the power of GWAS results. In total, we believe that the use of predictive traits with high accuracy facilitates the results of GWAS. In addition, there is no effect of weighted ssGBLUP on the ssGWAS results of this study under the condition of 50K chip data (detailed results not shown), which is consistent with other recent studies (Aguilar et al., 2019; Cesarani et al., 2021). One of the explanatory reasons for this result is that no SNP had a large effect on the traits studied.

4. Conclusions

This study explained the genetic background of NEI and its composition traits that can be used in dairy cattle breeding, and provides 16 key candidate genes influencing the genetic mechanism of NUE traits, which are mainly expressed in milk cell, mammary, and liver tissues. The NEI not only can reflect the six studied NUE-related traits but also can reflect the interaction between them, because only the NEI has a prominent peak at BTA14 explaining more than 0.50% of the total additive genetic variance. Furthermore, the NEI may be more representative of NUE because the genomic regions most associated with it are closer to genes in the nitrogen metabolism pathway.

5. Acknowledgments

The China Scholarship Council (Beijing) is acknowledged for funding the PhD project of Yansen Chen. The authors would like to thank the European Union’s Seventh Framework Program for funding the GplusE project under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. The authors thank Stefano Biffani give good suggestions for this work (IBBA-CNR National Research Council, Milano, Italia). The authors acknowledge the support of the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE; Namur, Belgium) and the use of the computation resources of the University of Liège – Gembloux Agro-Bio Tech (Gembloux, Belgium) provided by the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the F National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grants No. T.0095.19 (PDR “DEEPSELECT”) and J.0174.18 (CDR “PREDICT-2”). The authors have not stated any conflicts of interest.

6. Supplementary materials

The 6 supplemental files can be found on GitHub (HD_supplemetal_file.zip: https://github.com/Yansen0515/NEI_GWAS_POST_GWAS).

7. References

Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. Genet. Sel. Evol. 51:1–8. https://doi.org/10.1186/s12711-019-0469-3.

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 August 21, 2022. http://www.ainfo.inia.uy/digital/bitstream/item/15445/1/Aguilar-et-al.-2014.-WCGALP.pdf.

Aguirre-Villegas, H.A., M.A. Wattiauxand, R.A. Larson, L. Chase, S.D. Ranathunga, and M.D. Ruark. 2017. Dairy cow nitrogen efficiency. University of Wisconsin-Extension. Accessed May 28, 2022. https://www.virtualfarm.psu.edu/assets/uploads/content/DairyCap_DairyCowNitrogenEfficiency-Final-Final.pdf.

Ariyarathne, H.B.P.C., M. Correa-Luna, H.T. Blair, D.J. Garrick, and N. Lopez-Villalobos. 2021. Identification of genomic regions associated with concentrations of milk fat, protein, urea and efficiency of crude protein utilization in grazing dairy cows. Genes (Basel). 12:1–20. https://doi.org/10.3390/genes12030456.

Atashi, H., M. Salavati, J. De Koster, J. Ehrlich, M. Crowe, G. Opsomer, GplusE Consortium*, and M. Hostens. 2020. Genome-wide association for milk production and lactation curve parameters in Holstein dairy cows. J. Anim. Breed. Genet. 137:292–304. https://doi.org/10.1111/jbg.12442.

Bakhshalizadeh, S., S. Zerehdaran, and A. Javadmanesh. 2021. Meta-analysis of genome-wide association studies and gene networks analysis for milk production traits in Holstein cows. Livest. Sci. 250:1–14. https://doi.org/10.1016/j.livsci.2021.104605.

Bouwman, A.C., G.C.B. Schopen, H. Bovenhuis, M.H.P.W. Visker, and J.A.M. van Arendonk. 2010. Genome-wide scan to detect quantitative trait loci for milk urea nitrogen in Dutch Holstein-Friesian cows. J. Dairy Sci. 93:3310–3319. https://doi.org/10.3168/jds.2009-2829.

Brito, L.F., H.R. Oliveira, K. Houlahan, P.A.S. Fonseca, S. Lam, A.M. Butty, D.J. Seymour, G. Vargas, T.C.S. Chud, F.F. Silva, C.F. Baes, A. Cánovas, F. Miglior, and F.S. Schenkel. 2020. Genetic mechanisms underlying feed utilization and implementation of genomic selection for improved feed efficiency in dairy cattle. Can. J. Anim. Sci. 100:587–604. https://doi.org/10.1139/cjas-2019-0193.

Brunes, L.C., F. Baldi, F.B. Lopes, R.B. Lôbo, R. Espigolan, M.F.O. Costa, N.B. Stafuzza, and C.U. Magnabosco. 2021. Weighted single-step genome-wide association study and pathway analyses for feed efficiency traits in Nellore cattle. J. Anim. Breed. Genet. 138:23–44. https://doi.org/10.1111/jbg.12496.

Cai, Z., B. Guldbrandtsen, M.S. Lund, and G. Sahana. 2019. Dissecting closely linked association signals in combination with the mammalian phenotype database can identify candidate genes in dairy cattle. BMC Genet. 20:1–12. https://doi.org/10.1186/s12863-019-0717-0.

Calus, M. P., H. A. Mulder, and J. W. Bastiaansen. 2011. Identification of Mendelian inconsistencies between SNP and pedigree information of sibs. Genet. Sel. Evol. 43:1-12. https://doi.org/10.1186/1297-9686-43-34.

Calsamiglia, S., A. Ferret, C.K. Reynolds, N.B. Kristensen, and A.M. Van Vuuren. 2010. Strategies for optimizing nitrogen use by ruminants. Animal 4:1184–1196. https://doi.org/10.1017/S1751731110000911.

Cavani, L., M.B. Poindexter, C.D. Nelson, J.E.P. Santos, and F. Peñagaricano. 2022. Gene mapping, gene-set analysis, and genomic prediction of postpartum blood calcium in Holstein cows. J. Dairy Sci. 105:525–534. https://doi.org/10.3168/jds.2021-20872.

Cesarani, A., A. Garcia, J. Hidalgo, L. Degano, D. Vicario, N.P.P. Macciotta, and D. Lourenco. 2021. Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. J. Dairy Sci. 104:5719–5727. https://doi.org/10.3168/jds.2020-19838.

Chen, Y., H. Atashi, C. Grelet, S. Vanderick, H. Hu, and N. Gengler. 2022. Defining a nitrogen efficiency index in Holstein cows and assessing its potential impact on the breeding program of bulls. J. Dairy Sci. 105: 7575–7587. https://doi.org/10.3168/jds.2021-21681.

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.

de los Campos, G., A.I. Vazquez, S. Hsu, and L. Lello. 2018. Complex-Trait Prediction in the Era of Big Data. Trends Genet. 34:746–754. https://doi.org/10.1016/j.tig.2018.07.004

Elsik, C.G., R.L. Tellam, K.C. Worley, and The Bovine Genome Sequencing and Analysis Consortium*. 2009. The genome sequence of taurine cattle: A window to ruminant biology and evolution. Science, 324:522–528. https://doi.org/10.1126/science.1169588

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 9:1–9. https://doi.org/10.1093/gigascience/giaa149.

Fragomeni, B. de O., I. Misztal, D.L. Lourenco, I. Aguilar, R. Okimoto, and W.M. Muir. 2014. Changes in variance explained by top SNP windows over generations for three traits in broiler chicken. Front. Genet. 5:1–7. https://doi.org/10.3389/fgene.2014.00332.

Freebern, E., D.J.A. Santos, L. Fang, J. Jiang, K.L. Parker Gaddis, G.E. Liu, P.M. Vanraden, C. Maltecca, J.B. Cole, and L. Ma. 2020. GWAS and fine-mapping of livability and six disease traits in Holstein cattle. BMC Genomics 21:1–11. https://doi.org/10.1186/s12864-020-6461-z.

Gebreyesus, G., A.J. Buitenhuis, N.A. Poulsen, M.H.P.W. Visker, Q. Zhang, H.J.F. van Valenberg, D. Sun, and H. Bovenhuis. 2019. Combining multi-population datasets for joint genome-wide association and meta-analyses: The case of bovine milk fat composition traits. J. Dairy Sci. 102:11124–11141. https://doi.org/10.3168/jds.2019-16676.

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. 4435–4445. https://doi.org/10.3168/jds.2019-17910.

Honerlagen, H., H. Reyer, M. Oster, S. Ponsuksili, N. Trakooljul, B. Kuhla, N. Reinsch, and K. Wimmers. 2021. Identification of genomic regions influencing N-Metabolism and N-Excretion in lactating Holstein-Friesians. Front. Genet. 12:1–12. https://doi.org/10.3389/fgene.2021.699550.

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:D701–D710. https://doi.org/10.1093/nar/gky1084.

Kauffman, A.J., and N.R. St-Pierre. 2001. The relationship of milk urea nitrogen to urine nitrogen excretion in Holstein and Jersey cows. J. Dairy Sci. 84:2284–2294. https://doi.org/10.3168/jds.S0022-0302(01)74675-9.

Kolenda, M., and B. Sitkowska. 2021. The polymorphism in various milk protein genes in Polish Holstein-Friesian dairy cattle. Animals 11:1–8. https://doi.org/10.3390/ani11020389.

Li, B., L. Fang, D.J. Null, J.L. Hutchison, E.E. Connor, P.M. VanRaden, M.J. VandeHaar, R.J. Tempelman, K.A. Weigel, and J.B. Cole. 2019. High-density genome-wide association study for residual feed intake in Holstein dairy cattle. J. Dairy Sci. 102:11067–11080. https://doi.org/10.3168/jds.2019-16645.

Li, X., A.J. Buitenhuis, M.S. Lund, C. Li, D. Sun, Q. Zhang, N.A. Poulsen, and G. Su. 2015. Joint genome-wide association study for milk fatty acid traits in Chinese and Danish Holstein populations. J. Dairy Sci. 98:8152–8163. https://doi.org/10.3168/jds.2015-9383.

Liu, S., Y. Gao, O. Canela-xandri, S. Wang, Y. Yu, W. Cai, B. Li, R. Xiang, A.J. Chamberlain, E. Pairo-castineira, K.D. Mellow, K. Rawlik, C. Xia, Y. Yao, P. Navarro, D. Rocha, X. Li, Z. Yan, C. Li, B.D. Rosen, C.P. Van Tassell, P.M. Vanraden, S. Zhang, L. Ma, J.B. Cole, G.E. Liu, and A. Tenesa. 2022. A multi-tissue atlas of regulatory variants in cattle. Nat Genet. 54:1438–1447. https://doi.org/10.1038/s41588-022-01153-5.

Lourenco, D., A. Legarra, S. Tsuruta, Y. Masuda, I. Aguilar, and I. Misztal. 2020. Single-step genomic evaluations from theory to practice: using SNP chips and sequence data in blupf90. Genes (Basel). 11:1–32. https://doi.org/10.3390/genes11070790.

Minozzi, G., E.L. Nicolazzi, A. Stella, S. Biffani, R. Negrini, B. Lazzari, P. Ajmone-Marsan, and J.L. Williams. 2013. Genome wide analysis of fertility and production traits in Italian Holstein cattle. PLoS One 8:1–10. https://doi.org/10.1371/journal.pone.0080219.

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 May 28, 2022. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Nayeri, S., M. Sargolzaei, M.K. Abo-Ismail, N. May, S.P. Miller, F. Schenkel, S.S. Moore, and P. Stothard. 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. BMC Genet. 17:1–11. https://doi.org/10.1186/s12863-016-0386-1.

Nayeri, S., M. Sargolzaei, M.K. Abo-Ismail, S. Miller, F. Schenkel, S.S. Moore, and P. Stothard. 2017. Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle. J. Dairy Sci. 100:1246–1258. https://doi.org/10.3168/jds.2016-11770.

Olsen, H.G., T.M. Knutsen, A.M. Lewandowska-Sabat, H. Grove, T. Nome, M. Svendsen, M. Arnyasi, M. Sodeland, K.K. Sundsaasen, S.R. Dahl, B. Heringstad, H.H. Hansen, I. Olsaker, M.P. Kent, and S. Lien. 2016. Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the group-specific component (GC) gene in clinical mastitis and milk production. Genet. Sel. Evol. 48:1–16. https://doi.org/10.1186/s12711-016-0257-2.

Pegolo, S., N. MacH, Y. Ramayo-Caldas, S. Schiavon, G. Bittante, and A. Cecchinato. 2018. Integration of GWAS, pathway and network analyses reveals novel mechanistic insights into the synthesis of milk proteins in dairy cows. Sci. Rep. 8:1–15. https://doi.org/10.1038/s41598-017-18916-4.

Pinares-Patiño, C.S., G.C. Waghorn, A. Machmüller, B. Vlaming, G. Molano, A. Cavanagh, and H. Clark. 2007. Methane emissions and digestive physiology of non-lactating dairy cows fed pasture forage. Can. J. Anim. Sci. 87:601–613. https://doi.org/10.4141/CJAS06023.

Raven, L.A., B.G. Cocks, M.E. Goddard, J.E. Pryce, and B.J. Hayes. 2014. Genetic variants in mammary development, prolactin signalling and involution pathways explain considerable variation in bovine milk production and milk composition. Genet. Sel. Evol. 46:1–13. https://doi.org/10.1186/1297-9686-46-29.

Rijnkels, M., L. Elnitski, W. Miller, and J.M. Rosen. 2003. Multispecies comparative analysis of a mammalian-specific genomic domain encoding secretory proteins. Genomics 82:417–432. https://doi.org/10.1016/S0888-7543(03)00114-9.

Sanchez, M.P., A. Govignon-Gion, P. Croiseau, S. Fritz, C. Hozé, G. Miranda, P. Martin, A. Barbat-Leterrier, R. Letaïef, D. Rocha, M. Brochard, M. Boussaha, and D. Boichard. 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genet. Sel. Evol. 49:1–16. https://doi.org/10.1186/s12711-017-0344-z.

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

Serão, N.V.L., D. González-Peña, J.E. Beever, G.A. Bollero, B.R. Southey, D.B. Faulkner, and S.L. Rodriguez-Zas. 2013. Bivariate genome-wide association analysis of the growth and intake components of feed efficiency. PLoS One 8:1–11. https://doi.org/10.1371/journal.pone.0078530.

Soares, R.A.N., G. Vargas, T. Duffield, F. Schenkel, and E.J. Squires. 2021. Genome-wide association study and functional analyses for clinical and subclinical ketosis in Holstein cattle. J. Dairy Sci. 104:10076–10089. https://doi.org/10.3168/jds.2020-20101.

Strucken, E.M., R.H. Bortfeldt, D.J. De Koning, and G.A. Brockmann. 2012. Genome-wide associations for investigating time-dependent genetic effects for milk production traits in dairy cattle. Anim. Genet. 43:375–382. https://doi.org/10.1111/j.1365-2052.2011.02278.x.

Szklarczyk, D., A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, N.T. Doncheva, M. Legeay, T. Fang, P. Bork, L.J. Jensen, and C. von Mering. 2021. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49:D605–D612. https://doi.org/10.1093/nar/gkaa1074.

Taussat, S., M. Boussaha, Y. Ramayo-Caldas, P. Martin, E. Venot, G. Cantalapiedra-Hijar, C. Hozé, S. Fritz, and G. Renand. 2020. Gene networks for three feed efficiency criteria reveal shared and specific biological processes. Genet. Sel. Evol. 52:1–14. https://doi.org/10.1186/s12711-020-00585-z.

Uwizeye, A., I.J.M. de Boer, C.I. Opio, R.P.O. Schulte, A. Falcucci, G. Tempio, F. Teillard, F. Casu, M. Rulli, J.N. Galloway, A. Leip, J.W. Erisman, T.P. Robinson, H. Steinfeld, and P.J. Gerber. 2020. Nitrogen emissions along global livestock supply chains. Nat. Food 1:437–446. https://doi.org/10.1038/s43016-020-0113-y.

van den Berg, I., P.N. Ho, T.V. Nguyen, M. Haile-Mariam, I.M. MacLeod, P.R. Beatson, E. O’Connor, and J.E. Pryce. 2022a. GWAS and genomic prediction of milk urea nitrogen in Australian and New Zealand dairy cattle. Genet. Sel. Evol. 54:1–12. https://doi.org/10.1186/s12711-022-00707-9.

van den Berg, I., P.N. Ho, T. V Nguyen, M.H. Mariam, T.D.W. Luke, and J.E. Pryce. 2022b. Using mid - infrared spectroscopy to increase GWAS power to detect QTL associated with blood urea nitrogen. Genet. Sel. Evol. 54:1–8. https://doi.org/10.1186/s12711-022-00719-5

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

Veerkamp, R.F., A.C. Bouwman, C. Schrooten, and M.P.L. Calus. 2016. Genomic prediction using preselected DNA variants from a GWAS with whole-genome sequence data in Holstein-Friesian cattle. Genet. Sel. Evol. 48:1–14. https://doi.org/10.1186/s12711-016-0274-1.

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.

Wang, Z., X. Hou, B. Qu, J. Wang, X. Gao, and Q. Li. 2014. Pten regulates development and lactation in the mammary glands of dairy cows. PLoS One 9:1-13. https://doi.org/10.1371/journal.pone.0102118.

WHO and FAO (World Health Organization and Food and Agriculture Organization of the United Nations). 2011. Codex Alimentarius: Milk and Milk Products. 2nd ed. WHO FAO, Rome, Italy.

Wu, T., E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, and G. Yu. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation(China) 2:1-10. https://doi.org/10.1016/j.xinn.2021.100141.

Wiggans, G.R., T.S. Sonstegard, P.M. VanRaden, L.K. Matukumalli, R.D. Schnabel, J.F. Taylor, F.S. Schenkel, and C.P. van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. J. Dairy Sci. 92:3431–3436. https://doi.org/10.3168/jds.2008-1758.

Wilmot, H., J. Bormann, H. Soyeurt, X. Hubin, G. Glorieux, P. Mayeres, C. Bertozzi, and N. Gengler. 2022. Development of a genomic tool for breed assignment by comparison of different classification models: Application to three local cattle breeds. J. Anim. Breed. Genet. 139:40–61. https://doi.org/10.1111/jbg.12643.

Zhou, C., C. Li, W. Cai, S. Liu, H. Yin, S. Shi, Q. Zhang, and S. Zhang. 2019. Genome-wide association study for milk protein composition traits in a Chinese Holstein population using a single-step approach. Front. Genet. 10:1–17. https://doi.org/10.3389/fgene.2019.00072.

7

Chapter VII General discussion, perspectives, and conclusions

Foreword

The results of Chapters 2 to 6 supported the possibility of improving nitrogen use efficiency (NUE) and indirectly decreasing N pollution through genetic selection. The aims of this chapter are to compile results obtained throughout this thesis and to explore the possibility of adding an NUE proxy to the current breeding program in the Walloon Region of Belgium.

For these aims, the following points will be discussed: First, the biological background of the NUE proxies that were used in this thesis is discussed. Second, which NUE proxy to use in the breeding program is considered. Third, the consequences of adding the NUE proxy to the current breeding program are explained. Fourth, the possibility of improving the current optimal NUE proxy is proposed. Finally, a general conclusion will be drawn.

1. General discussion

The objective of this study was to explore the possibility of simultaneously obtaining both more economical and more respectful of the environment dairy cattle through genetic selection. As described in the introduction (Chapter 1), nitrogen (N) use efficiency (NUE) is one of the key traits for obtaining more economic and environmentally friendlier dairy cows. However, the NUE of dairy cows is very low, ranging from 15% to 44% (Cantalapiedra-Hijar et al., 2018; Grelet et al., 2020). Spanghero and Kowalski (2021) analyzed 86 experiments over the past 20 years and concluded that N retention may have been overestimated in N balance experiments, implying that the NUE of dairy cows may even be lower than reported. Recently, farmers in the Walloon Region of Belgium seek for dairy cows that are more suitable for grazing to reduce production costs (Lefèvre et al., 2022). Increased grazing time may potentially be associated with increased N pollution in air and water through urine and fecal N because of the N oversupply. At the same time, increasing grazing time will also encounter new challenges for farmers, such as feeding systems (Wales and Kolver, 2017), and stocking rates (Powell et al., 2010). The NUE in the farm gate level increased by 2.1% (from 20.8 to 22.9%) mainly because of reduced chemical N fertilizer inputs (Buckley et al., 2016). With environmental protection, researchers in New Zealand, a predominantly grazing country, have recently begun studies to reduce N pollution from cows (Beatson et al., 2019; Ariyarathne et al., 2019). Milk urea N (MUN), as a monitoring proxy for dairy cow protein nutrition, has been used to explore whether genetic selection could directly reduce N pollution and indirectly increase NUE in dairy cows (Marshall et al., 2020; Ariyarathne et al., 2021). Marshall et al. (2021, 2022) subsequently reported the effect of different MUN breeding values on N excretion by dairy cows under different forages. Another typical grazing country, Australia, has also started to research on breeding dairy cows for MUN (van den Berg et al., 2021, 2022).

Therefore, breeders in the Walloon Region need to find a suitable indicator to measure NUE and N pollution in dairy cows. As described in the introduction (Chapter 1), the NUE is a difficult and expensive trait to be directly recorded. So, the proxies of NUE were used in this thesis (Chapters 2 to 4). The biological background of NUE proxies needs to be first discussed when we use them. Then, which proxy is better used for the breeding selection program needs to be considered. Third, the consequences of adding the NUE proxies to the breeding program of the Walloon Region of Belgium are discussed. It is the effect of adding the optimal NUE proxy on the possible future add traits (e.g., methane, negative energy balance (NEB)).

1.1. Biological Background of proxies for Nitrogen Use Efficiency

The metabolic processes of N in cattle are described in the introduction section (Figure 1-1). Furthermore, the biological background of the NUE proxies included in this study is discussed, which will facilitate our more accurate selection and use of NUE proxies.

Milk urea concentration (MU), which is the traditional proxy for monitoring N pollution and protein efficiency in dairy cows, was found to be directly related to N excretion (Nousiainen et al., 2004; Burgos et al., 2007). The MU is derived from urea in plasma, and another destination of urea in plasma is to become urine urea in the kidney (Figure 1-1). The source of N in MU and urine is the same, which is the biological basis for MU to predict N pollution (Broderick and Clayton, 1997). The predicted NUE (PNUE), predicted N loss (PNL), and N efficiency index (NEI) are all referred to as novel proxies. Predicted NUE is calculated as total N output (Milk N) divided by total N intake (NINT) (Grelet et al., 2020). The disadvantages of PNUE are described in detail in Chapter 4, and we emphasize that one of the biological disadvantages of PNUE is that milk N contains MU. Milk N, while being increased, may potentially increase MU, which defeats our purpose of reducing N pollution. The PNL is another form of NUE proxy, which is NINT minus total output (Milk N). Its biological disadvantage is that it does not consider the N used in the maintenance and growth.

The NEI was newly proposed by us based on NINT, actual total output (milk true protein N, MTPN), and excretion fraction (MUN yield, MUNY) (Chapter 4). The NEI first controls the NINT without increasing or decreasing, which controls the total N metabolism by the bovine body from the source. The NINT is the most important influence on N pollution, and N excretion increases as NINT increases (Bougouin et al., 2022a). The MTPN is the output that we really need to increase, which is the absorption of N mainly from the plasma through the small intestine (Figure 1-1). The MUNY is used to reduce urinary N, and the biological background of its effect is the same as that of MU. The biological disadvantages of NEI are similar to those of PNL.

The metabolic processes of all NUE proxies involve multiple bovine organs. Organs directly related to NUE and N pollution in cows include the rumen, small intestine, liver, mammary, and kidney (Figure 1-1). The effect of the rumen on milk protein yield is through microorganisms and their metabolites (Xue et al., 2020). The small intestine is where N is absorbed by the body. The function of the liver for N metabolic is to convert ammonia into urea (Spek et al., 2013). The 16 key candidate genes related to NUE from results in Chapter 6 were also mainly expressed in liver and mammary tissues. Xie et al (2021) reported high feed efficiency (FE) (low residual feed intake) of cattle due to high amino acid utilization in the mammary. High FE also has high protein utilization (Liu and VandeHaar 2020), so we speculate high NUE cattle are also due to high amino acid usage in the mammary. The kidney can adjust the balance between MU and urea by reabsorption of urea (Spek et al., 2013).

1.2. Which Proxy is the Optimal Proxy of Nitrogen Use Efficiency in the Breeding Program

In addition to considering the biological background of the proxy, other factors need to be considered when selecting the optimal NUE proxy for use in a breeding program.

First, we need to consider whether the proxy can be acquired accurately and cheaply at a large scale. All proxies used in this study were predicted from milk mid-infrared (MIR) spectra (Chapters 2 to 4). Milk MIR predicted traits are inexpensive and rapid, so they can all be obtained on a large scale (Soyeurt et al., 2006). Milk MIR has now been routinely used for measuring milk protein, lactose, fat percentage, as well as MU worldwide (Gengler et al., 2016). So, MU is available to countries already undergoing Dairy Herd Improvement. Protein content in milk and MU are obtained directly from commercial instruments (e.g., FOSS). The prediction model parameters given in the FOSS company and literature are shown in Table 7-1. The relative error of the protein content and MU prediction model are lowest (< 1%), whereas the PNL is highest (23%). The MUN is accepted and used in genetic selection by researchers in New Zealand and Australia (Marshall et al., 2020; van den Berg et al., 2021). The model of NINT is considered to have good predictive ability relative to PNUE and PNL.

Table 7-1. Coefficient of determination (R^2_v) and relative error (%, root mean square error of validation data divided to mean) of prediction model for protein content (PC, g/L), milk urea concentration (MU, mg/dL), predicted nitrogen use efficiency (PNUE, %), predicted nitrogen losses (PNL, kg/d), and predicted nitrogen intake (NINT, kg/d)

Model	R^2_v	Relative error (%)	Reference
PC	/	< 1	FOSS
MU	/	< 1	FOSS
PNUE	0.68	14	Grelet et al., 2020
PNL	0.59	23	Grelet et al., 2020
NINT	0.71	15	Grelet et al., 2020

Second, the genetic correlation between NUE proxies needs to be considered. The genetic correlations between MU (representing urine N) and PNUE were low, ranging from -0.10 to 0.02 (shown in the 1.3 section of this Chapter). The reported genetic correlation between MUN and NUE was also very low (-0.07) (Tavernier et al., 2022).

This suggests that MU not be suitable for genetic selection as a proxy for NUE. So, we proposed a new NEI (Figure 4-1, Chapter 4) that considers both NUE and N pollution at the same time. The NEI was built by combining (genomic) estimated breeding values ((G)EBV) based on NUE-related traits: NINT, MTPN, and MUNY. Recently, the results of a large international collaboration also showed that predicting N excretion in dairy cows needs to include N intake and MUN to more accurately predict N pollution, and more region-specific models are needed (Bougouin et al., 2022a).

Based on the above discussion, NEI is considered in this thesis to be the current optimal proxy to increase NUE and reduce N pollution. In addition, we demonstrated that genomic prediction can be performed on NEI (Chapter 5).

1.3. Genetic Correlation between Milk Urea Concentration and Predicted Nitrogen Use Efficiency, Predicted Nitrogen Losses

These results added here are in support of the discussion in Section 1.2 and Chapter 4. The NEI in this study was established based on the result that MU has a low genetic correlation with PNUE (Milk N / NINT) and PNL (NINT - Milk N). The milk MIR data and model used in this work were the same as those used in the variance composition estimated part of Chapter 4. The filtering methods for MU, PNUE, and PNL were the same as in Chapter 2 and Chapter 3, respectively. Finally, the 143,595 records from 53,660 cows on 776 farms were used in this work. The pedigree related to the dataset comprised 132,057 animals. Besides, each feature was divided into primiparous and multiparous (including second to fifth parity) classes. Hereafter, the PNUE, PNL, and MU were identified as PNUE1, PNL1, MU1, PNUE2+, PNL2+, and MU2+. The variance components were estimated by Gibbs sampling in THRGIBBS1F90 (version 2.118), and posterior convergence was analyzed by POSTGIBBSF90 (version 3.14). Among them, the posterior means of variance were obtained using 100,000 samples after a burn-in of 20,000 samples. Using the mean of the posterior variance components, we calculated genetic and phenotype correlations. Approximate standard errors of all calculated parameters were obtained and implemented in the AIREMLF90 (version 1.144) program.

The correlations of the six traits were reported in Table 7-2. The correlations between PNUE and PNL were similar to our previous research results (Chapter 3). The genetic and phenotypic correlations of MU and PNUE, PNL ranged from –0.14 to 0.03, –0.16 to 0.08, respectively. However, the standard errors of these correlations between the above traits were relatively big. Therefore, we think that MU is not genetically related to PNUE, PNL. The reported genetic correlation between NUE and MU nitrogen (MUN) was also very low (-0.07) (Tavernier et al., 2022). We speculated that this may be due to the different metabolic mechanisms of MU and PNUE (Figure 1-1, Chapter 1). In addition, the MU1 and MU2 had a very high genetic correlation (0.99), which is the same as the results from Chapter 2.

Table 7-2. Genetic correlations (above the diagonal), and phenotype correlations (below the diagonal) among predicted nitrogen (N) use efficiency, predicted N losses and milk urea concentration of dairy cows in the primiparous (n = 44,321) and multiparous (from second to fifth parity, n = 99,275)

Trait[1]	PNUE1	PNL1	MU1	PNUE2+	PNL2+	MU2+
PNUE1		-0.77 ± 0.03	0.02 ± 0.06	0.73 ± 0.04	-0.61 ± 0.05	-0.02 ± 0.05
PNL1	-0.43 ± 0.01		-0.14 ± 0.06	-0.55 ± 0.04	0.81 ± 0.03	-0.09 ± 0.05
MU1	0.08 ± 0.01	-0.16 ± 0.01		-0.10 ± 0.05	-0.01 ± 0.05	0.99 ± 0.01
PNUE2+	0.19 ± 0.01	-0.11 ± 0.01	0.00 ± 0.01		-0.79 ± 0.03	-0.14 ± 0.04
PNL2+	-0.11 ± 0.01	0.18 ± 0.01	-0.02 ± 0.01	-0.47 ± 0.00[2]		0.03 ± 0.04
MU2+	0.01 ± 0.01	-0.04 ± 0.01	0.15 ± 0.01	0.06 ± 0.00	-0.10 ± 0.00	

[1] Trait: PNUE1 - Predicted N use efficiency in primiparous cows; PNL1 - Predicted N losses in primiparous cows; PNintake1 - Predicted N intake in primiparous cows; MU1 - Milk urea concentration in primiparous cows; PNUE2+ - Predicted N use efficiency in multiparous cows; PNL2+ - Predicted N losses in multiparous cows; PNintake2+ - Predicted N intake in multiparous cows; MU2+ - Milk urea concentration in multiparous cows.

[2]: 0.00 means less than 0.005

1.4. Consequences of adding the NUE proxies to the Current Breeding Program

The results for Chapters 2 to 5 all show that NUE proxies can be genetically selected, and the heritability ranges from 0.06 to 0.22. Moreover, the impact of the NUE proxy on other economically important traits needs to be considered when it is added to the breeding program, especially production and functional traits.

Different strength levels of genetic correlation between NUE proxies and other traits are shown in Chapters 2 to 4. The approximated genetic correlations (AGC) between MU and 11 other traits of interest were estimated throughout the whole lactation by the random regression model (Chapter 2). Our results indicate that MU selection has a positive effect on milk production (milk, fat, and protein yield) in early and late lactation (–0.25 to –0.10). However, MU selection had a negative effect on functional traits (e.g. longevity, female fertility) in early lactation (0.12 to 0.27) (Chapter 2). The AGC between novel proxies of NUE and other traits of interest were estimated only in the first 50 days in milk due to the limitations of the prediction models (Grelet et al., 2020). The positive AGCs between PNUE and production were got, and its greater in primiparous (0.24 to 0.47) than in multiparous cows (0.04 to 0.24) (Chapter 3). The

PNUE was negatively correlated with functional traits (–0.07 to –0.44), and the results were similar for primiparous and multiparous cows (Chapter 3). The NEI was positively genetically correlated with production traits (0.08 to 0.46) and negatively correlated with functional traits (–0.07 to –0.27) (Chapter 4). The genetic correlation between NEI and the global economic index is negative (–0.37) (Chapter 4), which shows that the current breeding program in the Walloon Region of Belgium is not conducive to NUE.

The results of this study indicated that the direct addition of NUE proxy in the breeding program had a beneficial effect on production, but not on functional traits. The actual measured NUE and milk yield have also been reported to have a positive genetic correlation (0.49) (Tavernier et al., 2022). Daughters of bulls selected for high NEI will have higher milk production and protein rates (Chapter 4). The results of Chapter 4 suggest that it is also feasible to select bulls with high NEI values and global economic index (the total index in the Walloon Region of Belgium). Therefore, adding appropriate economic values for NUE to the global economic index may not affect other traits, as González-Recio et al. (2020) did for greenhouse gases.

1.5. The Relationship between Nitrogen Efficiency Index and Methane

The main goal of this study is to investigate the possibility of selecting and breeding with more respect for the environment cows. Methane is considered as the major source of greenhouse gases emissions from cattle production systems that have a considerable contribution to climate change (Tricarico et al., 2022). Studies have shown that improving FE can reduce methane emissions from cattle (Manzanilla-Pech et al., 2021, 2022). As aforementioned high FE also has high protein use efficiency (Liu and VandeHaar 2020), however, the relationship between NUE and methane is unclear. Therefore, the genetic correlations between NEI and methane were estimated.

Vanlierde et al. (2021) recently developed a more robust predictive methane model based on reference data from seven countries. Atashi et al. (2022) performed a genetic analysis of the predicted methane from the above model. They used three methane expression units: one for daily predicted methane emission using the existing milk MIR equation (PME, g/d), the second is the predicted ratio of PME divided by the daily milk yield (MY, kg/d) recorded on the same test-day (PMI), and the third is LMI defined as the log-transformed of PMI. The distribution of LMI is more in line with the normal distribution, compared to PME and PMI. Therefore, the PME and LMI were chosen.

The AGC between NEI and PME and LMI were analyzed by the algorithm of Blanchard et al. (1983). The NEI and its reliability for bulls were obtained in Chapter

5 by using the whole dataset, whereas the breeding values of PME and LMI and their reliabilities were obtained using random regression test-day model (Atashi et al., 2022). A total of 575 bulls with NEI, PME, and LMI reliabilities of at least 0.50 were used in this analysis. The mean reliability values for NEI, PME, and LMI ranged from 0.72 to 0.89 (Table 7-3). The AGC between NEI and PME was –0.04, while that estimated between NEI and LMI was −0.43 (Table 7-3). This suggests that NEI is not genetically associated with PME but has a moderate negative genetic association with LMI. The FE, a trait similar to NUE, also showed different variations in genetic association with methane (Difford et al., 2020). Difford et al. (2020) explained that this could be caused by trait definition (residual feed intake in phenotypic and genetic levels), differences in environment and feed (grazing vs. Total mixed ration), etc. Although currently there is no procedure for researchers to change environmental and forage differences, research and communication can be increased to harmonize the definition of novel phenotypes. International cooperation can play a key role in the research and communication process through joint projects. This will promote the use of novel traits in breeding across countries. In conclusion, our results show that increasing NEI is beneficial in reducing the intensity of methane emissions from individual cows.

Table 7-3. Approximated genetic correlation (AGC, ± standard error) between nitrogen efficiency index (NEI) and PME, LMI evaluated in bulls[1] (n = 575)

	NEI	PME	LMI
Reliability			
Mean	0.72	0.89	0.89
Standard deviation	0.11	0.08	0.08
AGC with NEI	NA[2]	-0.04 ± 0.05	-0.43 ± 0.04

[1] PME: daily methane production using existing milk MIR equation (PME); LMI: log-transformed ratio of PME divided by the daily milk yield (kg/d) recorded on the same test-day

[2] NA: not applicable

1.6. The Relationship between Nitrogen Efficiency Index and Negative Energy Balance

One of the main goals of this study was to select cows with a higher NUE. Energy is one of the key factors affecting NUE (Rius et al., 2010). However, most high yield dairy cows in NEB are in early lactation because the energy intake of cows cannot meet their needs for milk production (Churakov et al., 2021). Breeding for NUE alone may enhance the NEB for dairy cows in early lactation. If cows are in the NEB state for too long, the risk of metabolic diseases (such as ketosis) increases (Zachut et al., 2020). The NEI establishment of this thesis (Chapter 4) was based on data from the first 50 days in milk, so we need to know the genetic relationship between NEI and NEB.

There is currently no efficient method to directly measure NEB in cows, but people usually predict it by indirect methods such as milk MIR spectra or biomarkers (milk C18:1 *cis*-9 and citrate) (Grelet et al., 2016; Ho et al., 2019; Churakov et al., 2021). When cows are in NEB, their body fat reserves will be mobilized, leading to an increase in C18:1 *cis*-9 and citrate in milk (Bjerre-Harpøth et al.,2012; Zachut et al., 2020). MIR can quickly detect and respond to changes in milk composition (Gengler et al., 2016). Although using milk MIR spectra to indirectly predict the NEB of dairy cows might have some sort of efficiency, the coefficient of determination of its prediction model varies among studies (from 0.48 to 0.77) (Ho et al., 2019; Smith et al., 2019). The C18:1 *cis*-9 and citrate can be accurately and quickly predicted using milk MIR. The determination coefficients for cross-validation of C18:1 *cis*-9 and citrate were 0.95 and 0.90 (Soyeurt et al., 2011; Grelet et al., 2014, 2016).

The C18:1 *cis*-9 and citrate using the above model (Grelet et al., 2014, 2016) based upon milk MIR in Chapter 5 were predicted. Then, a genetic analysis combining the NINT, MTPN, and MUNY in Chapter 5 was performed. The genetic analysis model and procedure were the same as in Chapter 5, except for the change to a 10-trait repeatability model. The NEI and its reliability were calculated in the same way as described in Chapter 5. The AGC between NEI and C18:1 *cis*-9 and citrate were calculated by using the algorithm of Blanchard et al. (1983). The 576 bulls with reliabilities at least 0.50 for all traits were used for this calculation.

Table 7-4. Approximated genetic correlation (AGC, ± standard error) between nitrogen efficiency index (NEI) and C18:1 *cis*-9, citrate based on EBV of bulls[1] (n = 576)

	NEI	C18:1 cis-9_1	Citrate1	C18:1 cis-9_2+	Citrate2+
Reliability					
Mean	0.72	0.72	0.65	0.86	0.70
Standard deviation	0.11	0.11	0.11	0.08	0.12
AGC with NEI	NA[2]	0.69 ± 0.05	0.73 ± 0.04	0.50 ± 0.05	0.65 ± 0.04

[1]: C18:1 *cis*-9_1 - C18:1*cis*-9 in primiparous cows; Citrate1 - Citrate in primiparous cows; C18:1 *cis*-9_2+ - C18:1 *cis*-9 in multiparous cows; Citrate2+ - Citrate in multiparous cows
[2]: NA - no application

The range of average reliability (standard deviation, SD) of NEI and C18:1 *cis*-9, citrate for the selected bulls was 0.65 (0.08) - 0.86 (0.12) (Table 7-4). The NEI and C18:1 *cis*-9, citrate had high positive AGC, from 0.50 to 0.73 (Table 7-4), which suggests that breeding for NUE alone may enhance energy issues in early lactation. Other efficiency traits are also consistent with our results, such as the genetic correlation between gross FE and energy balance in early lactation close to –1

(Spurlock et al., 2012), as well as the results of residual energy intake (Becker et al., 2021). The results demonstrate the need to also consider the NEB trait if NEI is considered for inclusion in the current breeding program.

1.7. The Relationship between Nitrogen Efficiency Index and Feed efficiency

The FE has been incorporated into dairy cattle breeding systems in some countries (e.g., Australia, the US, and Canada) as described in Chapter 4. The question is unanswered if NUE and FE are highly correlated, since crude protein accounts for approximately 18% of the dry matter in the diet. To address this question, we need to analyze the genetic relationship between NEI and FE traits. Dry matter intake (DMI) is one key indicator to assess the FE in dairy cows (Seymour, et al., 2019). The DMI can be predicted on a large scale by milk MIR (here called predicted DMI, PDMI), although its actual value is difficult and expensive to measure. The coefficient of determination and standard error in the cross-validation of the model for PDMI is 0.71 and 5.7 kg/d, respectively (Grelet et al., 2021).

The PDMI using the above model (Grelet et al.,2021) based on milk MIR in Chapter 5 was predicted. The PDMI are identified as PDMI1 for primiparous cows, and PDMI2+ for multiparous cows (second parity to fifth parity). Then, genetic analysis for PDMI was performed. The genetic analysis model and procedure were the same as in Chapter 5, except for the change to a two-trait repeatability model. The NEI and its reliability for bulls were from Chapter 5. The AGC between NEI and PDMI were calculated by using the algorithm of Blanchard et al. (1983). The 575 bulls with reliabilities at least 0.50 for all traits were used for this calculation.

Heritabilities of PDMI1 and PDMI2+ both were 0.19, which is similar to the real DMI in early lactation (Manzanilla Pech et al., 2014). The average reliability (SD) of NEI was the same as in section 1.5 of this Chapter. The average reliabilities (SD) of PDMI1 and PDMI2+ were 0.76 (0.10) and 0.80 (0.10), respectively. The AGC (standard error) between PDMI1, PDMI2+, and NEI were 0.32 (0.04) and 0.33 (0.04), respectively. The results suggest that increasing the DMI trait can potentially improve NEI. The DMI with feed conversion ratio in young bulls had a strong phenotypic positive correlation (0.86) (Callegaro et al., 2022). This might suggest that improving feed efficiency can partially improve NUE. Part of the explanation for this result is that NEI is established in early lactation, so DMI needs to be increased to compensate for the lack of metabolic energy. However, as expected, NUE was not highly genetically associated with FE traits. Combining the results of section 1.5 of this

Chapter, we suggest that when establishing efficiency traits in early lactation, the three aspects of nitrogen, energy, and the dry matter should be considered.

2. Perspectives

Different perspectives can be identified based on existing issues. First, there is a need for validation of NEI in populations with really measured N traits. The idea of high NUE is to maintain cattle N intake, increase production (milk N), and reduce N pollution (urinary and fecal N). However, because of the lack of measurement data on N pollution, the hypothesis that improving NEI will reduce urinary N through reducing MUNY depends on the validity of their linear relationship. A suggested validation test would first select high and low NEI from local cattle according to NEI, and then carry out the recording of real N traits to verify the response. It would also allow for further optimization of the NEI based on the data measured in this test. As stated in the discussion section in Chapter 4, the addition of live weight or body condition score (BCS) to the existing NEI is a relatively straightforward strategy to improve it. In the Walloon Region of Belgium, BCS of cows is taken in routine with conformation recording, but mostly once and in the first lactation only. At present, our group is trying to build a model for predicting BCS using milk MIR (N. Gengler, personal communication) that could be added to direct BCS data. Obtaining live weight through automatic scales is possible. Also, efforts exist to predict live weight using routine data in farms and milk MIR (Tedde et al., 2021). The latest NUE reported by Tavernier et al. (2022) considered N fixed in the muscle and conceptus, which brings the NUE closer to the true value. If urinary and fecal N values are obtained, we can then calculate a more accurate index weight directly from biology. A prediction equation for total excreted N based on N intake was recently reported (Bougouin et al., 2022a), which will facilitate our work. On the other hand, we built the NEI based on (G)EBV of six traits which are considered genetic levels. We can also build new NEI with phenotypes of six traits, which is considering genetic and environmental interactions.

Second, we should validate the genomic predictions of NEI on a larger dataset, as discussed in Chapter 5. This could be done by increasing the reference population, which is one of the common ways to improve the accuracy of genome predictions. Another way to improve the accuracy of genomic prediction is to use a set of single nucleotide polymorphisms (SNP) associated with the predicted traits (Xiang et al., 2021). We can generate a set of SNPs related to NUE traits based on the results of Chapter 6. Next, explore the accuracy of genome prediction based on the selected SNP set and compare it with the results of Chapter 6. Tahir et al. (2022) recently showed that SNP sets generated from multi-omics data can improve the accuracy of genome prediction for heifer fertility traits. In addition, using some new algorithms may also improve the accuracy of genomic prediction, such as genomics omics BLUP (GOBLUP) (Legarra and Christensen, 2022), mixed effect neural networks (NN-MM) (Zhao et al., 2022), or machine learning (Wang et al., 2022).

Third, the current predictors of NUE are limited to early lactation because of the prediction model limitations. We believe it is necessary to develop predictors that can be used throughout the whole lactation for the following two reasons. First of all, the NUE is constantly changing throughout the lactation period (Tavernier et al., 2022). On the other hand, the efficiency of crude protein utilization, which is similar to NUE, has a large variation in genetic parameters throughout the lactation period (Ariyarathne et al., 2021). Ongoing local work (C. Grelet, personal communication) and work done by other groups show that this should be possible. Shi et al. (2022) developed models to predict NUE in mid and late lactation. The coefficients of determination of the prediction equations (shown in Chapter 1) in validation were similar to those reported by Grelet et al. (2020). This demonstrates that milk MIR can predict NUE throughout lactation in dairy cows. One of the future goals is to update the model with data collected into mid and late lactation, which will allow us to predict NUE-related traits (e.g., NINT) throughout lactation. Gengler (2022) illustrated how equation building could be organized by an innovative collaborative open consortium to build a new NUE model that predicts the whole lactation. Then, the genetic correlation between NEI and other traits can be analyzed throughout lactation, as described in Chapter 2. This allows analysis of the potential effects of NEI on the functional traits of cows throughout lactation. The results presented in Chapter 2 showed that the genetic association of MU with longevity, female fertility, and calving ease was greater in early than late lactation.

Adding NEI to the existing broader breeding goal (V€G) for dairy cattle in the Walloon Region of Belgium is the fourth issue. As we said in Chapter 5, the Walloon genetic evaluation system is composed of different sub-indexes. It would use the NEI as part of a broader efficiency sub-index included in a future V€G (Vanderick et al., 2022). The broader efficiency sub-index may consist of NEI, methane, energy, and DMI. Breeding for improved NUE in cattle could be initiated by this approach. Richardson et al. (2020, 2021, 2022) recently reported the inclusion of methane in Australian cattle breeding goals giving us many indications for future work.

This work focused on dairy cattle, but not on non-lactating and beef cattle. Bougouin et al. (2022b) integrated an international database to show that the NUE of beef cattle is only 22% showing a potential fifth issue. Therefore, it is also interesting to consider increasing permanently the NUE of non-lactating beef through genetic selection. This requires more work. First, it is important to clarify the definition of NUE for non-lactating and beef cattle, such as the new NUE = (N intake - urine N - fecal N) / N intake. These three phenotypes are difficult to record, and it is an issue that researchers have been trying to breakthrough. Angelidis et al. (2019) attempted to build predictive models for urine and fecal N, but the variables they used were difficult to record routinely on commercial farms. Bougouin et al. (2022b) also established urine and fecal N prediction models, but models with sufficient accuracy all require N intake for prediction. Therefore, the development of an inexpensive detection method for N intake of individual cattle should be the focus of future work. The introduction of advanced technologies such as sensors or image recognition may solve this difficulty.

3. Conclusions

This study showed that it is possible to develop genetic selection for both more economical and more respectful of the environment cattle. The h^2 of proxies of NUE were ranging from 0.09 to 0.22. The genetic selection for proxies (MU, PNUE, PNL, and NEI) of NUE will result in a positively correlated response in production traits, while resulting in a negatively correlated response for functional traits in early lactation (Chapters 2 to 4). The MU may not be suitable for genetic selection as a proxy for NUE. The newly established NEI can be genetically selected as the index for NUE and N pollution. In addition, increasing NEI in early lactation favors a reduction in the intensity of methane emissions and increases dry matter intake but is detrimental to energy balance. However, the genetic correlation between NEI and other traits may be different in the whole lactation according to the results from Chapter 2. The NEI can perform genomic prediction was verified (Chapter 5). Finally, the genomic background of NEI and its composition traits were investigated (Chapter 6). The 16 identified genes and their high-expression tissue locations can be used for future studies on the genetic mechanisms of NUE.

Furthermore, research from the genome to the NUE phenotype is just beginning. With the release of the pangenome for global cattle (e.g., Talenti et al., 2022; Zhou et al., 2022) and cattleGTEx (Liu et al., 2022), researchers can better understand the genetic expression process of the NUE and other phenotypes.

4. References

Angelidis, A., L. Crompton, T. Misselbrook, T. Yan, C.K. Reynolds, and S. Stergiadis. 2019. Evaluation and prediction of nitrogen use efficiency and outputs in faeces and urine in beef cattle. Agric. Ecosyst. Environ. 280:1–15. https://doi.org/10.1016/j.agee.2019.04.013.

Ariyarathne, H., H.T. Blair, and D.J. Garrick. 2019. Estimation of heritabilities for milk urea and efficiency of crude protein utilization by day of lactation in a New Zealand dairy cow herd. New Zeal. J. Anim. Sci. Prod. 79:183–187.

Ariyarathne, H.B.P.C., M. Correa-Luna, H. Blair, D. Garrick, and N. Lopez-Villalobos. 2021. Can nitrogen excretion of dairy cows be reduced by genetic selection for low milk urea nitrogen concentration? Animals 11:1–18. https://doi.org/10.3390/ani11030737.

Atashi, H, H. Wilmot, S. Vanderick, X. Hubin, H. Soyeurt, and N. Gengler. 2022. G enome-wide association study for mid-infrared methane predictions in Walloon dairy cows. Proceedings of the 12th World Congress on Genetics Applied to Li vestock Production, Rotterdam, The Netherlands. Accessed July 12, 2022. https: //www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/69_003.pdf.

Beatson, P.R., S. Meier, N.G. Cullen, and H. Eding. 2019. Genetic variation in milk urea nitrogen concentration of dairy cattle and its implications for reducing urinary nitrogen excretion. Animal 13:2164–2171. https://doi.org/10.1017/S1751731119000235.

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.

Bougouin, A., A. Hristov, J. Dijkstra, M.J. Aguerre, S. Ahvenjärvi, C. Arndt, A. Bannink, A.R. Bayat, C. Benchaar, T. Boland, W.E. Brown, L.A. Crompton, F. Dehareng, I. Dufrasne, M. Eugène, E. Froidmont, S. van Gastelen, P.C. Garnsworthy, A. Halmemies-Beauchet-Filleau, S. Herremans, P. Huhtanen, M. Johansen, A. Kidane, M. Kreuzer, B. Kuhla, F. Lessire, P. Lund, E.M.K. Minnée, C. Muñoz, M. Niu, P. Nozière, D. Pacheco, E. Prestløkken, C.K. Reynolds, A. Schwarm, J.W. Spek, M. Terranova, A. Vanhatalo, M.A. Wattiaux, M.R. Weisbjerg, D.R. Yáñez-Ruiz, Z. Yu, and E. Kebreab. 2022a. Prediction of nitrogen excretion from data on dairy cows fed a wide range of diets compiled in an intercontinental database: A meta-analysis. J. Dairy Sci. 105:7462–7481. https://doi.org/10.3168/jds.2021-20885.

Bougouin, A., A. Hristov, D. Zanetti, S.C. V Filho, L.N. Rennó, A.C.B. Menezes, J.M. Silva Jr, H.M. Alhadas, L.D.S. Mariz, L.F. Prados, K.A. Beauchemin, T. McAllister, W.Z. Yang, K.M. Koenig, K. Goossens, T. Yan, P. Noziere, A. Jonker, and E. Kebreab. 2022b. Nitrogen excretion from beef cattle fed a wide range of diets compiled in an intercontinental dataset: a meta-analysis. J. Anim. Sci. 100:1–12. https://doi.org/10.1093/jas/skac150.

Broderick, G.A., and M.K. Clayton. 1997. A Statistical Evaluation of Animal and Nutritional Factors Influencing Concentrations of Milk Urea Nitrogen. J. Dairy Sci. 80:2964–2971. https://doi.org/10.3168/jds.s0022-0302(97)76262-3.

Buckley, C., D.P. Wall, B. Moran, S. O'Neill, and P.N.C. Murphy. 2016. Farm gate level nitrogen balance and use efficiency changes post implementation of the EU Nitrates Directive. Nutr. Cycl. Agroecosystems 104:1–13. https://doi.org/10.1007/s10705-015-9753-y.

Burgos, S.A., J.G. Fadel, and E.J. DePeters. 2007. Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: Relation of milk urea nitrogen to urine urea nitrogen excretion. J. Dairy Sci. 90:5499–5508. https://doi.org/10.3168/jds.2007-0299.

Callegaro, S., G. Niero, M. Penasa, R. Finocchiaro, G. Invernizzi, and M. Cassandro. 2022. Greenhouse gas emissions, dry matter intake and feed efficiency of young Holstein bulls. Ital. J. Anim. Sci. 21:870–877. https://doi.org/10.1080/1828051X.2022.2071178.

Cantalapiedra-Hijar, G., R.J. Dewhurst, L. Cheng, A.R.J. Cabrita, A.J.M. Fonseca, P. Nozière, D. Makowski, H. Fouillet, and I. Ortigues-Marty. 2018. Nitrogen isotopic fractionation as a biomarker for nitrogen use efficiency in ruminants: A meta-analysis. Animal 12:1827–1837. https://doi.org/10.1017/S1751731117003391.

Chen, Y., C. Grelet, S. Vanderick, and N. Gengler. 2021a. Relationship between proxies of nitrogen use efficiency for dairy cows in early lactation. Accessed April 9, 2022. https://orbi.uliege.be/bitstream/2268/263087/1/EAAP21_71_YansenCHEN.pdf.

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 15:1–10. https://doi.org/10.1016/j.animal.2021.100253.

Difford, G.F., P. Løvendahl, R.F. Veerkamp, H. Bovenhuis, M.H.P.W. Visker, J. Lassen, and Y. de Haas. 2020. Can greenhouse gases in breath be used to genetically improve feed efficiency of dairy cows? J. Dairy Sci. 103:2442–2459. https://doi.org/10.3168/jds.2019-16966.

FOSS. CombiFoss^{TM} 7- Herd-improvement and payment testing in one integrated unit. Accessed Dec. 14, 2022. https://www.fossanalytics.com/en/products/combifoss-7.

Gengler, N. 2022. Collaboration is imperial or how to address this critical issue when developping novel traits. The 73rd Annual Meeting of the European Federation of Animal Science, Porto, Portugal. Oral.

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, and 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.

González-Recio, O., J. López-Paredes, L. Ouatahar, N. Charfeddine, E. Ugarte, R. Alenda, and J.A. Jiménez-Montero. 2020. Mitigation of greenhouse gases in dairy cattle via genetic selection: 2. Incorporating methane emissions into the breeding goal. J. Dairy Sci. 103:7210–7221. https://doi.org/10.3168/jds.2019-17598.

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., 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.

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.

Lefèvre, A., V. Decruyenaere, and E. Froidmont. 2022. Transition towards more robust cows on Walloon dairy farms for a better use of forage resources. Dair’Innov 2022 congress, Namur, Belgium.

Legarra, A., and O.F. Christensen. 2022. Genomic evaluation methods to include intermediate correlated features such as high-throughput or omics phenotypes. JDS Commun. In press. https://doi.org/10.3168/jdsc.2022-0276.

Liu, S., Y. Gao, O. Canela-xandri, S. Wang, Y. Yu, W. Cai, B. Li, R. Xiang, A.J. Chamberlain, E. Pairo-castineira, K.D. Mellow, K. Rawlik, C. Xia, Y. Yao, P. Navarro, D. Rocha, X. Li, Z. Yan, C. Li, B.D. Rosen, C.P. Van Tassell, P.M. Vanraden, S. Zhang, L. Ma, J.B. Cole, G.E. Liu, and A. Tenesa. 2022. A multi-tissue atlas of regulatory variants in cattle. Nat Genet. 54:1438–1447. https://doi.org/10.1038/s41588-022-01153-5.

Manzanilla-Pech, C.I.V., P. L⊘vendahl, D. Mansan Gordo, G.F. Difford, J.E. Pryce, F. Schenkel, S. Wegmann, F. Miglior, T.C. Chud, P.J. Moate, S.R.O. Williams, C.M. Richardson, P. Stothard, and J. Lassen. 2021. Breeding for reduced methane emission and feed-efficient Holstein cows: An international response. J. Dairy Sci. 104:8983–9001. https://doi.org/10.3168/jds.2020-19889.

Manzanilla-Pech, C.I.V., R.B. Stephansen, G.F. Difford, P. Løvendahl, and J. Lassen. 2022. Selecting for Feed Efficient Cows Will Help to Reduce Methane Gas Emissions. Front. Genet. 13:1–10. https://doi.org/10.3389/fgene.2022.885932.

Manzanilla Pech, C.I.V., R.F. Veerkamp, M.P.L. Calus, R. Zom, A. van Knegsel, J.E. Pryce, and Y. De Haas. 2014. Genetic parameters across lactation for feed intake, fat- and protein-corrected milk, and liveweight in first-parity Holstein cattle. J. Dairy Sci. 97:5851–5862. https://doi.org/10.3168/jds.2014-8165.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2020. Grazing dairy cows with low milk urea nitrogen breeding values excrete less urinary urea nitrogen. Sci. Total Environ. 739:1–8. https://doi.org/10.1016/j.scitotenv.2020.139994.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2021. Nitrogen balance of dairy cows divergent for milk urea nitrogen consuming either plantain or perennial ryegrass. Animals 11:1–18. https://doi.org/10.3390/ani11082464.

Marshall, C.J., M.R. Beck, K. Garrett, G.K. Barrell, O. Al-Marashdeh, and P. Gregorini. 2022. Urine and fecal excretion patterns of dairy cows divergent for milk urea nitrogen breeding values consuming either a plantain or ryegrass diet. J. Dairy Sci. 105:4218–4236. https://doi.org/10.3168/jds.2021-21490.

Nousiainen, J., K.J. Shingfield, and P. Huhtanen. 2004. Evaluation of milk urea nitrogen as a diagnostic of protein feeding. J. Dairy Sci. 87:386–398. https://doi.org/10.3168/jds.S0022-0302(04)73178-1.

Powell, J.M., C.J.P. Gourley, C.A. Rotz, and D.M. Weaver. 2010. Nitrogen use efficiency: A potential performance indicator and policy tool for dairy farms. Environ. Sci. Policy 13:217–228. https://doi.org/10.1016/j.envsci.2010.03.007.

Richardson, C.M., P.R. Amer, F.S. Hely, I. van den Berg, and J.E. Pryce. 2021. Estimating methane coefficients to predict the environmental impact of traits in the Australian dairy breeding program. J. Dairy Sci. 104:10979–10990. https://doi.org/10.3168/jds.2021-20348.

Richardson, C.M., P.R. Amer, C. Quinton, J. Crowley, F.S. Hely, I. van den Berg, and J.E. Pryce. 2022. Reducing greenhouse gas emissions through genetic selection in the Australian dairy industry. J. Dairy Sci. 105:4272–4288. https://doi.org/10.3168/jds.2021-21277.

Richardson, C.M., C.F. Baes, P.R. Amer, C. Quinton, P. Martin, V.R. Osborne, J.E. Pryce, and F. Miglior. 2020. Determining the economic value of daily dry matter intake and associated methane emissions in dairy cattle. Animal 14:171–179. https://doi.org/10.1017/S175173111900154X.

Rius, A.G., M.L. McGilliard, C.A. Umberger, and M.D. Hanigan. 2010. Interactions of energy and predicted metabolizable protein in determining nitrogen efficiency in the lactating dairy cow. J. Dairy Sci. 93:2034–2043. https://doi.org/10.3168/jds.2008-1777.

Seymour, D.J., A. Cánovas, C.F. Baes, T.C.S. Chud, V.R. Osborne, J.P. Cant, L.F. Brito, B. Gredler-Grandl, R. Finocchiaro, R.F. Veerkamp, Y. de Haas, and F. Miglior. 2019. Invited review: Determination of large-scale individual dry matter intake phenotypes in dairy cattle. J. Dairy Sci. 102:7655–7663. https://doi.org/10.3168/jds.2019-16454.

Shi, R., W. Lou, B. Ducro, A.v.d. Linden, H.A. Mulder, S.J. Oosting , S. Li, and Y. Wang. 2022. Predicting nitrogen use efficiency, nitrogen loss and dry matter intake of individual dairy cows in late lactation by including mid-infrared spectra of milk samples. Research Square. https://doi.org/10.21203/rs.3.rs-1722648/v1.

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.

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.

Spanghero, M., and Z.M. Kowalski. 2021. Updating analysis of nitrogen balance experiments in dairy cows. J. Dairy Sci. 104:7725–7737. https://doi.org/10.3168/jds.2020-19656.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, and A. Bannink. 2013. A review of factors influencing milk urea concentration and its relationship with urinary urea excretion in lactating dairy cattle. J. Agric. Sci. 151:407–423. https://doi.org/10.1017/s0021859612000561.

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

Tahir, M.S., L.R. Porto-Neto, A. Reverter, B.S. Olasege, M.R. Sajid, K.B. Wockner, A.W.L. Tan, and M.R.S. Fortes. 2022. Utility of multi-omics data to inform genomic prediction of heifer fertility traits. J. Anim. Sci. 100:1–10. https://doi.org/10.1093/jas/skac340.

Talenti, A., J. Powell, J.D. Hemmink, E.A.J. Cook, D. Wragg, S. Jayaraman, E. Paxton, C. Ezeasor, E.T. Obishakin, E.R. Agusi, A. Tijjani, K. Marshall, A. Fisch, B.R. Ferreira, A. Qasim, U. Chaudhry, P. Wiener, P. Toye, L.J. Morrison, T. Connelley, and J.G.D. Prendergast. 2022. A cattle graph genome incorporating global breed diversity. Nat. Commun. 13:1–14. https://doi.org/10.1038/s41467-022-28605-0.

Tavernier, E., I.C. Gormley, L. Delaby, S. McParland, M. O'Donovan1 and D.P. Berry. 2022. Genetic variation exists in nitrogen use efficiency for lactating cows in grass-grazed dairy systems. Proceedings of the 12th World Congress on Genetics Applied to Livestock Production, Rotterdam, The Netherlands. Accessed July 12, 2022. https://www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/14_008.pdf.

Tedde, A., P.N. Ho, J.E. Pryce, D. Hailemariam, Z. Wang, G. Plastow, N. Gengler, Y. Brostaux, E. Froidmont, C. Bertozzi, M.A. Crowe, I. Dufrasne, GplusE

Consortium, and H. Soyeurt. 2021. Validation of Dairy Cow Bodyweight Prediction Using Traits Easily Recorded by Dairy Herd Improvement Organizations Selection Algorithms. Animals. 11:1–15. https://doi.org/10.3390/ani11051288.

Tricarico, J.M., Y. de Haas, A.N. Hristov, E. Kebreab, T. Kurt, F. Mitloehner, and D. Pitta. 2022. Symposium review: Development of a funding program to support research on enteric methane mitigation from ruminants. J. Dairy Sci. 105:8535–8542. https://doi.org/10.3168/jds.2021-21397.

van den Berg, I., P.N. Ho, M. Haile-Mariam, P.R. Beatson, E. O'Connor, and J.E. Pryce. 2021. Genetic parameters of blood urea nitrogen and milk urea nitrogen concentration in dairy cattle managed in pasture-based production systems of New Zealand and Australia. Anim. Prod. Sci. 61:1801–1810. https://doi.org/10.1071/AN21049.

van den Berg, I., P.N. Ho, T.V. Nguyen, M. Haile-Mariam, I.M. MacLeod, P.R. Beatson, E. O'Connor, and J.E. Pryce. 2022. GWAS and genomic prediction of milk urea nitrogen in Australian and New Zealand dairy cattle. Genet. Sel. Evol. 54:1–12. https://doi.org/10.1186/s12711-022-00707-9.

Vanderick, S., R. R. Mota, K. Wijnrocx, and N. Gengler. 2022. Description of the genetic evaluation systems used in the Walloon Region of Belgium. Accessed Nov. 21, 2022. http://www.elinfo.be/indexEN.html.

Vanlierde, A., F. Dehareng, N. Gengler, E. Froidmont, S. McParland, M. Kreuzer, M. Bell, P. Lund, C. Martin, B. Kuhla, and H. Soyeurt. 2021. Improving robustness and accuracy of predicted daily methane emissions of dairy cows using milk mid-infrared spectra. J. Sci. Food Agric. 101:3394–3403. https://doi.org/10.1002/jsfa.10969.

Wales, W.J., and E.S. Kolver. 2017. Challenges of feeding dairy cows in Australia and New Zealand. Anim. Prod. Sci. 57:1366–1383. https://doi.org/10.1071/AN16828.

Wang, X., S. Shi, G. Wang, W. Luo, X. Wei, A. Qiu, F. Luo, and X. Ding. 2022. Using machine learning to improve the accuracy of genomic prediction of reproduction traits in pigs. J. Anim. Sci. Biotechnol. 13:1–12. https://doi.org/10.1186/s40104-022-00708-0.

Xiang, R., I.M. MacLeod, H.D. Daetwyler, G. de Jong, E. O'Connor, C. Schrooten, A.J. Chamberlain, and M.E. Goddard. 2021. Genome-wide fine-mapping identifies pleiotropic and functional variants that predict many traits across global cattle populations. Nat. Commun. 12:1–13. https://doi.org/10.1038/s41467-021-21001-0.

Xie, Y., C. Miao, Y. Lu, H. Sun, and J. Liu. 2021. Nitrogen metabolism and mammary gland amino acid utilization in lactating dairy cows with different residual feed intake. Anim. Biosci. 34:1600–1606. https://doi.org/10.5713/AB.20.0821.

Xue, M.Y., H.Z. Sun, X.H. Wu, J.X. Liu, and L.L. Guan. 2020. Multi-omics reveals that the rumen microbiome and its metabolome together with the host metabolome contribute to individualized dairy cow performance. Microbiome 8:1–19. https://doi.org/10.1186/s40168-020-00819-8.

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.

Zhao, T., J. Zeng, and H. Cheng. 2022. Extend mixed models to multilayer neural networks for genomic prediction including intermediate omics data. Genetics 221:1–10. https://doi.org/10.1093/genetics/iyac034.

Zhou, Y., L. Yang, X. Han, J. Han, Y. Hu, F. Li, H. Xia, L. Peng, C. Boschiero, B.D. Rosen, D.M. Bickhart, S. Zhang, A. Guo, C.P.V. Tassell, T.P.L. Smith, L. Yang, and G.E. Liu. 2022. Assembly of a pangenome for global cattle reveals missing sequences and novel structural variations, providing new insights into their diversity and evolutionary history. Genome Res. 32:1585–1601. https://www.genome.org/cgi/doi/10.1101/gr.276550.122.

8

Chapter VIII Appendix

1. Peer-reviewed scientific publications

Chen, Y., H. Atashi, C. Grelet, R.R. Mota, S. Vanderick, H., Hu, GplusE Consortium, and N. Gengler. 2022. Genome-wide association study and functional annotation analyses for nitrogen efficiency index and its composition traits in dairy cattle. J. Dairy Sci. (accepted)

Chen, Y., H. Atashi, C. Grelet, S. Vanderick, H., Hu, and N. Gengler. 2022. Defining a nitrogen efficiency index in Holstein cows and assessing its potential impact on the breeding program of bulls. J. Dairy Sci. 105: 7575–7587. https://doi.org/10.3168/jds.2021-21681.

Chen, Y., H. Atashi, S. Vanderick, R.R. Mota, H. Soyeurt, H. Hammami, and N. Gengler. 2021. Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows. J. Dairy Sci. 104:12741–12755. https://doi.org/10.3168/jds.2021-20659.

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. 104:4412–4423. https://doi.org/10.3168/jds.2020-18849.

2. Contributions to international conferences

Chen, Y., S. Vanderick, C. Grelet, H. Hu., and N. Gengler. 2022. Relationship between proxies of energy states and nitrogen use efficiency for Holstein cows in early lactation. Proceedings of the 12th World Congress on Genetics Applied to Livestock Production, Rotterdam, The Netherlands. (poster)

Chen, Y., and N. Gengler. 2021. Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in Holstein. International Frontiers Forum on Animal Genetics, Breeding, and Reproduction. Wuhan, China. (oral, online)

Chen, Y., S. Vanderick, C. Grelet, and N. Gengler. 2021. Relationship between proxies of nitrogen use efficiency for dairy cows in early lactation. The 72nd Annual Meeting of the European Federation of Animal Science. Davos, Switzerland. (oral)