From the Institute of Genome Biology

Research Institute for Farm Animal Biology (FBN) in Dummerstorf

and Professorship of Animal Breeding and Genetics

Faculty of Agricultural and Environmental Sciences

University of Rostock

Host genetic and rumen microbial determinants of nitrogen (N) utilisation and N excretion in lactating Holsteins

Cumulative Dissertation

To the acquisition of the academic degree

Doctor of Agriculture (doctor agriculturae)

Faculty of Agricultural and Environmental Sciences

University of Rostock

Submitted by MSc. Hanne Honerlagen

born on 20.02.1995 in Kiel

Rostock, 2022

https://doi.org/10.18453/rosdok_id00004473

Gutachter:

Prof. Dr. Klaus Wimmers

Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät

Forschungsinstitut für Nutztierbiologie, Institut für Genombiologie

Prof. Dr. Christa Kühn

Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät

Forschungsinstitut für Nutztierbiologie, Institut für Genombiologie

Prof. Dr. Jörn Bennewitz

Universität Hohenheim, Fakultät der Agrarwissenschaften und Nutztierwissenschaft

Prof. Dr. Jens Tetens

Georg-August-Universität Göttingen, Department für Nutztierwissenschaften

Jahr der Einreichung: 2022

Jahr der Verteidigung: 2023

The ones that stand always by my side

Mama, Papa, Caro, Nele & Eiki

List of contents

List of abbreviations

AA	amino acid
ATP	adenosine triphosphate
BU	blood urea concentration
CP	crude protein
DAG	differential abundant (microbial) genus
DEG	differential expressed gene
dGBV	direct genomic breeding value
DIM	days in milk
DM	dry matter
DMI	dry matter intake
EBV	estimated breeding value
FPR	fat to protein ratio (milk)
GBV_{HMU}	direct genomic breeding value for high milk urea phenotype
GBV_{LMU}	direct genomic breeding value for low milk urea phenotype
GBV_{MED}	direct genomic breeding value for medium milk urea phenotype
GBVMU	genomic breeding value for milk urea
gEBV	genomic enhanced breeding value
GWAS	genome-wide association study
h^2	heritability
h^2_B	heritability of a specific microbial abundance (microbial phenotype)
HMU	high milk urea phenotype
HMUg	predisposition for high milk urea phenotype
LMU	low milk urea phenotype
LMUg	predisposition for low milk urea phenotype
LP	low (crude) protein diet
ME	metabolisable energy
MIR	mid-infrared
MU	milk urea concentration
N	nitrogen
NGS	next-generation sequencing
NH_3	ammonia
NH^{4+}	ammonium-ion
NP	normal (crude) protein diet
NPN	non-protein N
NUE	nitrogen use efficiency
OTU	operational taxonomic unit
RNB	ruminal nitrogen balance
RZG	total breeding merit index (Germany)
RZM	relative (weight of) milk performance complex (part of the German breeding goal)
SNP	single nucleotide polymorphism
UDP	rumen undegraded protein
UU	urinary urea concentration
vst	variance-stabilizing transformation

List of figures

List of tables

1 Abstract (Kurzfassung)

Enhancing the nitrogen (N) utilisation efficiency (NUE) of dairy cows by breeding selection was considered relevant for environmental footprints, feeding costs and animal health. Due to the difficult data obtainment of NUE, the cow-individual milk urea (MU) value was suggested as an indicator trait for breeding purposes. However, the complexity of N metabolism in ruminants has prevented a uniform understanding of MU's accuracy to picture NUE yet. It became evident that a deeper knowledge on the underlying principles of cow-individual N utilisation and N excretion is indispensable before MU can be seriously considered for breeding strategies on enhanced NUE.

Considering the key function of the rumen microbiome in the N metabolism it was hypothesised that cow-individual NUE is determined by the entire "holobiont", thus by the host genome and the rumen microbiome. In this context four study approaches were conducted, which focused on the host genome–trait, the host genome–rumen microbiota and the entire host genome–rumen microbiota–trait axis, aiming for a holistic exploration of cow-individual N utilisation and N excretion as a pre-step for future breeding strategies.

The studies were based on data material of two different cow populations. The 1st and 4th studies were conducted on data of 371 genotyped, lactating Holsteins, which were sampled once each for rumen fluid, milk, urine and faeces on a practice- operating dairy farm. For 358 cows, a genomic breeding value for MU (GBVMU) was estimated. The 2nd and 3rd studies were based on a population of 20 non-pregnant lactating Holsteins, which were grouped due to their predisposition for high or low MU values (HMUg *vs.* LMUg). The cows were housed on an experimental farm and fed with either a normal or a low CP diet (NP *vs.* LP), in a 4x5 balanced design (NPxHMUg, NPxLMUg, LPxHMUg, LPxLMUg; n= 5 cows each group). The 20 cows were slaughtered after two weeks and sampled for rumen tissue and rumen fluids directly after slaughter.

Initially, the host genome–trait axis was explored by a genome-wide association study (GWAS) for nine N traits in milk and urine. Potential candidate genes were identified, namely *GJA1*, *RXFP1*, *FRY1* (MU), *SH3D19* (MU yield), *RCAN2, CLIC5, ENPP4*, *ENPP5* (urinary urea concentration (UU)), *ELF2* and *SLC7A11* (minor N fractions in milk), and *ITPR2*, *MYBPC1*, *STIM2*, *SGCD*, *SLC6A2*, *TMCC2* and *MFSD4A* (specific non-urea-N metabolites in urine).

The 2nd and 3rd study focused on the host genome–rumen microbiota axis. The analysis of 16S rRNA amplicon sequencing data (microbial data) and holistic transcriptome profiling (host- gene expression) of the rumen epithelia identified significantly differential abundances of 10 microbial genera (DAG) and 28 gene transcripts (DEG) which distinguished HMUg and LMUg cows' rumen profiles. Greater abundances of the ureolytic genus *Succinivibrionaceae_UCG-002* and *unclassified Ruminococcaceae*

were identified in LMUg animals, whereas HMUg cows had enhanced occurrences of the *Butyrivibrio* genus. Differential expression analysis revealed genes of the bovine Major Histocompatibility Complex (*BOLA* genes) as well as *MX1*, *ISG15* and *PRSS2* displaying candidates of MU predisposition that were further attributed to enhanced immune system activity in LMUg cows. The analysis of a potential microbial–host interplay revealed 157 significantly correlated microbe–gene pairs and two pronounced microbe–transcript clusters, which jointly distinguished the rumen profiles of HMUg and LMUg cows. Strikingly positive correlations of *BOLA-DRA* transcripts with abundances of *Roseburia* and the *Lachnospiraceae* family might constitute particularly prominent microbial–host interplays of MU predisposition. The reduction of feed N was followed by 18 DAG in HMUg and 19 DAG in LMUg, depicting pronounced interest on *Shuttleworthia*, which displayed controversial adaption in HMUg and LMUg cows. Lowering feed N further elicited massive downregulation of immune response and energy metabolism pathways in LMUg cows.

The 4th study focused on the exploration of the entire host genome–rumen microbiota–trait axis. Therefore, the large cow population was grouped for GBVMU into two extreme groups (GBV_{HMU}, GBV_{LMU}, n=59) and one medium group (GBV_{MED} n=299). The analysis of 16SrRNA microbial data by contrasting the extreme groups identified a microbial signature of 24 genera, which potentially inferred from GBVMU selection. In accordance with the results of the 2nd and 3rd study, the microbial signature uncovered higher abundance of the ureolytic *Succinivibrionaceae_UCG-002* in GBV_{LMU} cows, whereas GBV_{HMU} hosted higher abundance of hyper ammonia producing bacteria. The abundances of the microbial signature were further investigated for correlations to proxies of NUE in milk, urine and faeces in the entire cow population (n=358). Three genera of the *Lachnospiraceae* family revealed significant correlations by their ruminal abundances to MU values, proposing them as considerable players in the host genome– rumen microbiota–MU axis. The significant correlations of *Prevotellaceae UCG-003, Anaerovibrio, Blautia* and *Butyrivibrio* abundances with MU, milk nitrogen and the N concentration in faeces (FaecN) suggested their contribution to genetically determined N utilisation in Holstein cows. Moreover, noticeable positive correlations between *WCHB1-41 ge*, *CAG-352* and *Bacteroidales BS11 gut group* abundances to FeacN suggested a contribution of these genera to genetically determined N losses due to non-assimilated N.

The study results were further discussed in regard of MU as a potential indicator trait for NUE. Significantly lower MU and UU values in GBV_{LMU} compared to GBV_{HMU} cows, similarities of DAG identified in both cow populations and a possible correction of MU records by cow-individual N intake data in the near future let assume potential for MU data being utilised for future breeding programs.

Kurzfassung

Die züchterische Selektion von Milchkühen für eine erhöhte Stickstoff (N) Nutzungseffizienz (NUE) hat ökologische, ökonomische und tiergesundheitliche Relevanz. Aufgrund der schwierigen Datenerhebung von NUE-Phänotypen wurde der Milchharnstoffwert (MU) als potenzielles Indikatormerkmal für die Zucht vorgeschlagen, jedoch wurde dessen Aussagekraft über NUE kontrovers diskutiert.

Die Schlüsselfunktion des Pansenmikrobioms in dem N- Metabolismus von Wiederkäuern lässt vermuten, dass die kuhindividuelle NUE durch den gesamten „Holobionten", also durch das Wirtsgenom und das Pansenmikrobiom, determiniert wird. Die vorliegende Dissertation untersucht diese „holobiontische" Determinierung der kuhindividuellen N- Nutzung und N- Ausscheidung, mit Fokus auf MU als potenzielles Zuchtmerkmal. Dafür wurden vier Studien durchgeführt, welche die Wirtsgenom–Phänotyp, die Wirtsgenom–Pansenmikrobiom, sowie die gesamte Wirtsgenom–Pansenmikrobiom–Phänotyp Achse untersuchen.

Für die Studien wurden zwei Kuhpopulationen beprobt. Die eine Population (Studie 1 und 4) bestand aus 371 genotypisierten, laktierenden Holsteinkühen, welchen auf einem Praxisbetrieb einmalig Pansensaft-, Urin-, Kot- und Milch entnommen wurde. Für 358 Tiere war ein genomischer Zuchtwert für MU (GBVMU) verfügbar. Die zweite Kuhpopulation (Studie 2 und 3) bestand aus 20 laktierenden Holsteinkühen, welche basierend auf ihrer Veranlagung für hohe oder niedrige MU-Werte (HMUg *vs.* LMUg) paarweise in einer experimentellen Stallanlage gehalten und mit einer Normal- oder einer Niedrigrohproteinration (NP *vs.* LP) gefüttert wurden (NPxHMUg, NPxLMUg, LPxHMUg, LPxLMUg; n= 5 Kühe pro Gruppe). Nach zweiwöchiger Fütterungsphase und anschließender Schlachtung wurden von den 20 Tieren Pansensaft- und Pansengewebsproben gewonnen.

In der ersten Studie wurde die Wirtsgenom–Phänotyp Achse mithilfe einer genomweiten Assoziationsstudie (GWAS) für N- Metaboliten in Milch und Urin untersucht. Als potenzielle Kandidatengene wurden *GJA1*, *RXFP1*, *FRY1* (MU), *SH3D19* (MU Menge), *RCAN2, CLIC5, ENPP4*, *ENPP5* (Urinharnstoffkonzentration (UU)), *ELF2*, *SLC7A11* (kleine N- Metaboliten in der Milch) und *ITPR2*, *MYBPC1*, *STIM2*, *SGCD*, *SLC6A2*, *TMCC2* und *MFSD4A* (Nicht-Harnstoff-N Metaboliten im Urin) identifiziert.

Mithilfe einer 16S rRNA Amplicon Sequenzierung des Pansenmikrobioms und einer holistischen Transkriptomanalyse des Pansengewebes wurde in der zweiten und dritten Studie die Wirtsgenom–Pansenmikobiom Achse untersucht. Die Pansenprofile von HMUg und LMUg Tieren zeigten signifikant unterschiedliche Vorkommen von zehn Mikrobengattungen (DAG) und 28 Gentranskripten (DEG). Im Vergleich zu HMUg Tieren wiesen LMUg Kühe höhere Vorkommen der harnstoffspaltenden Mikrobengattungen *Succinivibrionaceae_UCG-002* und von bislang unklassifizierten *Ruminococcaceae*

Gattungen auf, zeigten gleichzeitig aber geringere Vorkommen der *Butyrivibrio*. Die Identifikation höherer Transkriptabundanzen von *BOLA*-Genen, *MX1*, *ISG15* und *PRSS2* wiesen auf eine signifikant höhere, ruminale Immunaktivität in LMUg, verglichen zu HMUg Tieren hin. Eine Korrelationsanalyse zwischen Mikroben- und Transkriptabundanzen identifizierte 157 signifikant korrelierende Mikroben–Genpaare, welche in zwei Clustern konträr korrelierten. Hoch signifikant positive Korrelationen zeigte unter anderem die Transkriptabundanz von *BOLA-DRA* mit der *Roseburia* Gattung und der *Lachnospiraceae* Familie. Die Fütterung der LP-Ration führte zu einer ähnlichen Adaption der Mikroben in beiden Kuhgruppen, jedoch zeigte die *Shuttleworthia* Gattung eine konträre Adaption in LMUg verglichen zu HMUg Tieren. Infolge der Futter- N Reduktion wiesen LMUg Tiere zudem ein signifikant geringeres Vorkommen von Gentranskripten des Immunsystems, sowie eine Inaktivierung einiger Energiestoffwechsel- Genpfade auf.

Für die vierte Studie wurde die große Kuhpopulation mithilfe der GBVMU Werte eingeteilt (GBV_{HMU} n=29; GBV_{LMU}, n=30; GBV_{MED}, n=299), um den Gesamtkomplex aus Wirtsgenom–Pansenmikrobiom–Phänotyp zu untersuchen. Die Ergebnisse der Mikrobenanalyse identifizierte eine Mikrobensignatur von 24 Gattungen, deren unterschiedliche Vorkommen die Pansenprofile der Extremgruppen (GBV_{HMU} *vs.* GBV_{LMU}) teilweise auftrennten. Analog zu dem HMUg *vs.* LMUg Vergleich (Studie 2 und 3), zeigte auch der GBV_{LMU} *vs.* GBV_{HMU} Vergleich ein signifikant höheres Vorkommen der harnstoffspaltenden Mikrobengattung *Succinivibrionaceae_UCG-002* und geringere Vorkommen von Hyper-Ammonium produzierenden Bakterien in GBV_{LMU} Kühen. Eine Korrelationsanalyse zwischen dem Vorkommen der Mikrobensignatur und NUE assoziierten Phänotypen in der gesamten Kuhpopulation (n= 358 Kühe) zeigte drei Gattungen der *Lachnospiraceae* als potenziell wichtige Kandidaten der Wirtsgenom–Pansenmikroben–MU Phänotyp Achse auf. Die Korrelationen der Vorkommen von *Prevotellaceae UCG-003, Anaerovibrio, Blautia* und *Butyrivibrio zu* MU, dem Milch-N Gehalt und der fäkalen N-Konzentration (FaecN) ließen einen Zusammenhang dieser Mikroben in der wirtsabhängigen N-Nutzung und N- Ausscheidung vermuten. Moderate Korrelationen zwischen dem Vorkommen der *WCHB1-41 ge*, *CAG-352* und *Bacteroidales BS11 gut group* mit FaecN Werten deuteten zudem auf einen Zusammenhang dieser Gattungen zu der N- Ausscheidung von nicht-assimilierten N hin.

Die Studienergebnisse wurden abschließend im Zuge der Evaluierung von MU als potenzielles Zuchtmerkmal diskutiert. Signifikant geringere MU und UU-Werte in GBV_{LMU} verglichen zu GBV_{HMU} Kühen, kongruente Ergebnisse der Mikrobenanalysen in beiden Kuhpopulationen, sowie Fortschritte in der Datenerhebung zur kuhindividuellen Futteraufnahme sprechen für die Nutzung von MU als Indikatormerkmal in der Milchviehzucht.

2 Motivation

High yielding dairy cows consume about 0.5 kg nitrogen (N)/day via dietary crude protein (CP) intake. In symbiosis with their rumen microbes, the dairy cow partly utilises this N intake for milk and body protein synthesis but excretes the main part via urine and faeces. On average N excretions from dairy cows account for approximately 130 kg N/per cow and year (Martens *et al.*, 2021). Urinary and faecal-N excretions contribute substantially to climate change and groundwater pollution by the volatilisation of nitrous oxide and nitrate, and therefore attain political relevance (Uwizeye *et al.*, 2020). Considering that dietary CP accounts on average for 15% of the entire feed ration, N excretions constitute a remarkable nutrient loss with major economic relevance for the farmer (Tan *et al.*, 2021). Furthermore, the cow transforms absorbed, but not utilised N into urea to prevent ammonia intoxification, which is an energy consuming process (Munyaneza *et al.*, 2017). Thus, N excretions are crucial from a socio-ecological, economic and animal's health perspective.

Several studies have subjected the dietary N utilisation and N excretions of dairy cows, aiming to enhance NUE. Nutritional and physiological approaches aimed for the optimized dietary N supply at herd level, whereas genetic studies focused on inter– individual variance in nitrogen use efficiency (NUE) for breeding purpose. Breeding selection for enhanced NUE would equalise the dietary N requirement within a herd and accordingly decrease dietary N demands and losses from dairy farms as a long-term stable and environmental independent tool. However, up to now, the accurate collection of NUE phenotypes has not been feasible at a large scale. The resulting lack in genetic parameter estimation as well as the negative correlations of efficiency traits to further breeding goals have prevented the implementation of NUE into breeding programs yet (Pryce *et al.*, 2014, Chen *et al.*, 2021b). A deeper knowledge on the fundamental principles of NUE predisposition represents an indispensable step before the exploitation of inter-individual NUE variability can be seriously considered for breeding strategies.

3 State of the art

The following sections provide an overview on the recent knowledge of cow-individual NUE, the current breeding practice and considerations for NUE as being implemented into dairy cows' breeding program.

3.1 N metabolism in dairy cows

N is part of the functional group in amino acids (AA) and provides a nutritional value for mammals in the organically bound form of proteins, AA and (di)-peptides. Whereas monogastrics are dependent on a specific dietary AA pattern, ruminants are enabled to receive an autarkic AA provision from dietary protein and non-protein-N (NPN) metabolites due to the protein synthesis of a complex microbiome, which is allocated in the complex forestomach system of ruminants (Riemeier, 2004, Von Engelhardt *et al.*, 2015). The ruminant's forestomach constitutes three compartments (reticulum, rumen, omasum), which account for 80-90% of the weight of all gastric compartments and for about 50% of the entire gastrointestinal tract in the adult. The forestomach system fills the entire left abdominal cavity of the ruminant and has a capacity of about 100 liters in the adult cow (Von Engelhardt *et al.*, 2015). The rumen acts as a pre-gastric, anaerobic fermentation chamber in which the rumen microbiota is located (Cholewińska *et al.*, 2020). The rumen microbiota constitute bacteria, archaea and fungi, with bacteria accounting for the most abundant group with a density of 10^9–10^{11} ml^{-1} in the rumen fluid (Von Engelhardt *et al.*, 2015). The composition of the rumen bacteria was estimated to contain more than 5000 species – compared to approximately 1400 species in carnivores, omnivores and human beings – with the phyla *Firmicutes, Bacteroidetes, Proteobacteria, Fibrobacteres,* and to a small amount *Tenericutes* and *Actinobacteria* as most abundant members (representing 67% of all species) (Khafipour *et al.*, 2016).

The cow acquires N via the dietary CP. CP is calculated by multiplying the dietary N content with the default nitrogen-to-protein conversion factor 6.25 (Jones factor). Thus, CP encompasses all dietary N compounds, *i.e.,* protein-N, (di-) peptide-N, AA-N, and other N compounds, such as urea, nitrate and uric acid (Mariotti *et al.*, 2008). A small part of CP passes the rumen as rumen undegraded protein (UDP), and is either digested by host enzymes in the small intestine and subsequently absorbed, or it is excreted via faeces (Figure 1). The major amount of dietary N is provided to the rumen microbes, which incorporate parts of this ruminal N influx into microbial protein (= microbial growth). In the course of microbial protein synthesis, the rumen microbiota transform C-chains and N metabolites into valuable AA and thus enable the host – in comparison to monogastrics — the usage of all dietary N compounds (dietary protein-N and non-protein-N (NPN)) as well as an autarkic protein provision (Virtanen, 1966). The resulting microbial protein pattern is subsequently degraded to AA and (di-) peptides in the cow's intestine and absorbed via the intestinal surface. (Bergen *et al.*, 1968). Depending

on the ruminal energy supply, the microbial protein synthesis was estimated to yield approximately 10g protein/MJ^{-1} metabolisable energy (ME) (Von Engelhardt *et al.*, 2015). For dairy cows, the microbial protein comprises on average 40-80% of bioavailable N in the intestine and thus represents the most relevant AA source of the host (Riemeier, 2004). The part of ruminal N influx, which is not directly incorporated into microbial protein or excreted via faeces, is microbially converted to ammonia (NH_3), ammonium-ions (NH^{4+}), peptides and AA (Bryant, 1970, Cholewińska *et al.*, 2020). Peptides and AA are further utilised in the rumen by the microbes or absorbed via intestinal surface by the host. NH_3 and NH^{4+} either are utilised by NH_3-assimilating microbes for microbial growth or discharge the rumen via rumen epithelial absorption. The rumen epithelium converts the absorbed NH_3 and NH^{4+} via glutamate dehydrogenase to α-Ketoglutarate. The resultant molecule glutamate (and partly glutamine) is subsequently transported with the blood stream to the liver. In the hepatocytes, N is primarily detached from its carrier and further detoxified by the urea cycle. The detoxification of ammonia is an energy-consuming process, which requires four adenosine triphosphate (ATP) per yielded urea molecule (Von Engelhardt *et al.*, 2015). Urea is further secreted back into the blood cycle (blood urea) from which it is eliminated via urine (urinary urea) and milk (milk urea), or secreted back into the rumen across the rumen villi and via saliva (Von Engelhardt *et al.*, 2015).

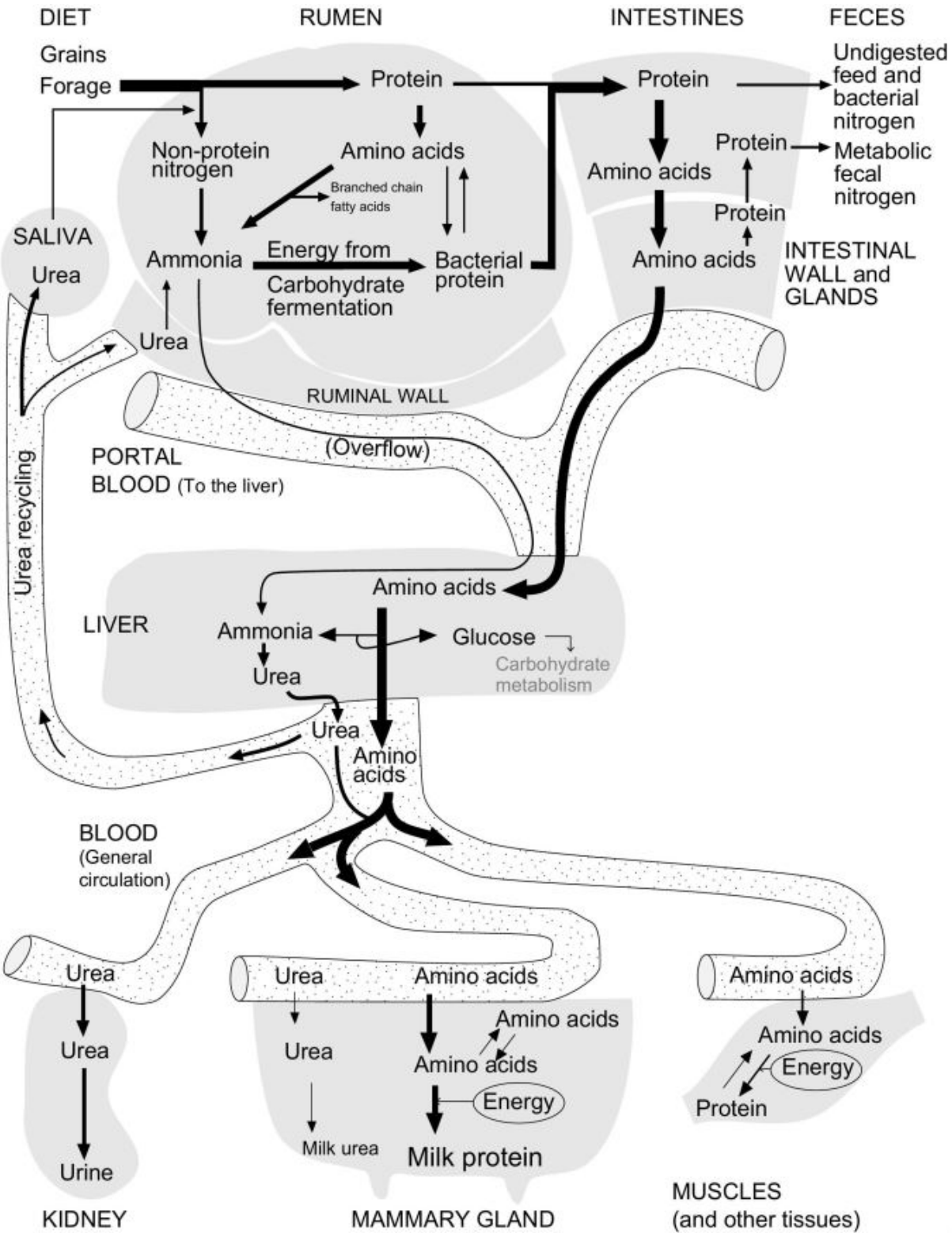

Figure 1. Overview of protein metabolism in dairy cows* (Wattiaux, 1998).

*published at https://kb.wisc.edu/dairynutrient/page.php?id=52745; access: 08.11.2022; author permitted usage of the original figure.

Urea is a polar molecule, which passes cell membranes alongside concentration gradients by diffusion (Von Engelhardt *et al.*, 2015, Lavery and Ferris, 2021). The urea concentrations of blood (BU), urine (UU) and milk (MU) were found to be closely related (Broderick and Clayton, 1997, Burgos *et al.*, 2007). The reflux of urea into the rumen increases as a consequence of dietary CP reduction due to the physiological mechanism of ruminants, which enables a decreased renal urea clearance rate and an increased renal urea reabsorption under N scarcity in order to retain N (Von Engelhardt *et al.*, 2015, Lavery and Ferris, 2021).

The recycled urea is converted by epithelium-bound, ureolytic bacteria in the rumen to CO_2 and NH_3. The latter, in turn, can be utilised by NH_3-assimilating microbes for microbial protein synthesis (Jin *et al.*, 2018). However, the rate of microbial protein synthesis depends on various factors, such as ruminal pH, dietary energy content, properties of the CP and the composition of the microbial community (Bryant, 1970).

3.2 N use efficiency (NUE) of dairy cows

The N use efficiency (NUE) of lactating cows is defined as the ratio of N in milk to dietary N intake (Calsamiglia *et al.*, 2010). In the context of environmental relevance, NUE for dairy cows was specified as the ratio of milk protein-N to feed N and NH_3-N losses (Groenestein *et al.*, 2019). Moreover, Houlahan *et al.* (2021) interpreted the term ”most efficient cow” as the cow that “produces the same amount of milk and milk solids while consuming less feed, but remaining healthy and fertile”. In regard of the published literature, NUE of dairy cows will address to the most commonly utilised definition: NUE= N in milk (g)/dietary N intake (g) (Calsamiglia *et al.*, 2010, Lavery and Ferris, 2021) in the following sections and will refer to the ratio of N in product (g) to dietary N intake (g) for other species.

3.2.1 NUE ranges in dairy cows

Although the ontogeny of ruminants allowed these species theoretically the utilisation of all dietary N compounds, the NUE of ruminants is – compared to monogastrics – low. Average NUEs of cattle, goats and sheep were documented with 25%, 15.6% and 16% respectively, whereas pigs and rats utilise dietary N with 35% and 60% each (Kohn *et al.*, 2005). Average NUE in dairy cows was almost uniformly documented with 25%, but wide ranges between cow-individual NUEs were observed (Calsamiglia *et al.*, 2010). Excluding sampling and measurement errors, factors that determine differences in NUEs can be categorised into dietary and animal characteristics, the latter being further separated into non-genetic (*i.e.,* parity, lactation stage) and genetic factors (Ferreira *et al.*, 2021).

3.2.1.1 NUE ranges in dairy cows: dietary effect

Dietary CP concentration and the calculated ruminal N balance (RNB) of the feed ration were documented as the most important determinants of NUE variability (Guliński *et al.*, 2016). The RNB is an indicator, which is used by the German feeding recommendation system to estimate the adequacy of dietary N supply to the rumen microbes (Kand and Dickhoefer, 2021). RNBs are calculated for single feed compounds of a diet by the formula: $RNB = \frac{CP - host\ available\ CP}{6.25}$ (Flachowsky *et al.*, 2001). Host available CP represents that amount of CP, which is supposed to be present in forms of AA and (di-) peptides in the intestine, thus mainly origins from microbial protein and to a small amount from UDP (see above: chapter 3.1). The dependency of microbial protein synthesis on the dietary energy and N supply, and the recycling of NPN into the rumen (see above, chapter 3.1), led to the calculation of negative RNBs for energy feed compounds (*e.g.*: whole plant silage of barley, wheat, rye or triticale= -3 g/kg DM (LfL, 2013); maize silage= -7g/kg DM (Tiefenthaller, 2009)), and to positive RNBs for dietary protein sources (*e.g.*: broad bean= +16g/kg DM; soy bean= +32 g/kg DM; rape seed= +22g/kg DM (LfL, 2013); grass silage 1st cut= +5g/kg DM (Tiefenthaller, 2009)). Based on the entire feed ration composition, the RNBs of single feed compounds are added up and result either in a positive or negative RNB per kg dry matter intake (DMI). In general, feed rations with negative RNBs aim for the increased ruminal utilisation of recycled NPN by the microbes (Martens *et al.*, 2021) and can be achieved by proportionally higher energy to lower protein feed compounds in the feed ration (Riemeier, 2004).

The reduction of dietary CP concentration and the reduction of dietary RNBs were consistently proven to increase NUE (*e.g.:* (Colmenero and Broderick, 2006, Huhtanen and Hristov, 2009, Lee *et al.*, 2011, Giallongo *et al.*, 2015, Mutsvangwa *et al.*, 2016, Schiavon *et al.*, 2016, Martens *et al.*, 2021, Müller *et al.*, 2021)). Reviewing the material of five studies with different N input levels (289g to 1360g N per cow and day), NUE ranges from 16% to 36% were published (Powell *et al.*, 2010). Individual NUEs of 40 lactating Holsteins increased from 29.6% to 40.3% when dietary CP was reduced from 19.6% to 13.5% of dry matter (DM) (Colmenero and Broderick, 2006). Similarly, NUEs of 20 lactating Holsteins increased (30.7% to 36%) among dietary CP reduction (15.9 to 13.8% of DM) in another study (Müller *et al.*, 2021). A decrease of the RNB from -10g/day to -40g/day increased NUEs from 35% to 37%, which was observed by cow-individual data of 11,375 Holsteins stalled in an experimental farm (Martens *et al.*, 2021). A potential effect of CP quality on NUE levels and ranges was explored by comparing the NUEs of 287 dairy cows, which were fed, with either grass based or maize based diets (Calsamiglia *et al.*, 2010). NUEs comprised 21.0% to 32.0% for the grass based and 22.0% to 32.8% for the maize-based diets, assuming no effect of CP quality on cow-individual NUE.

3.2.1.2 NUE ranges in dairy cows: non-genetic cow effects

Addressing to non-genetic cow characteristics, the stage of lactation was assumed to affect NUE, mainly due to the negative energy balance in early lactation, which is – among other effects – accompanied by protein catabolism (Lavery and Ferris, 2021). The accuracy of a prediction model for NUE increased when stage of lactation was included with the categories <49 days in milk (DIM), 50 – 99 DIM, 100 to 189 DIM and >190 DIM (Lahart *et al.*, 2019). Moreover, in a dataset of 129 early lactating Holsteins the parity correlated significantly positive with NUE (parities: 1st, 2nd, 3rd; r^2= 0.26, *p*- value < 0.001) (Grelet *et al.*, 2020).

3.2.1.3 NUE ranges between dairy cows: genetic cow aspects

Cow-individual NUE is supposed to be influenced by genetics (Stamer *et al.*, 2009, Grelet *et al.*, 2020, Chen *et al.*, 2021b, Lavery and Ferris, 2021). However, a reliable heritability (h^2) estimation for NUE requires large datasets of cow-individual NUE phenotypes. Due to the difficulty in obtaining data of cow-individual DMI and of accurate dietary compositions on farm, individual NUE phenotypes are still limited to sample sizes on experimental farms (Lavery and Ferris, 2021). In a recent approach, genetic parameters were estimated for a prediction model of NUE (pNUE), which was established on milk mid-infrared (MIR) spectral data (Grelet *et al.*, 2020, Chen *et al.*, 2021b). The model was developed in a dataset of 129 Holsteins in early lactation, which were stalled in three different countries and repeatedly measured for NUE and MIR data. The prediction accuracy of the model was stated with 82% (Grelet *et al.*, 2020). Heritabilities for pNUE were subsequently estimated on the basis of 46,163 milk samples from 21,462 cows and comprised 0.13 and 0.12 for 1st and later parities (Chen *et al.*, 2021b). Concerns of the model were mentioned, such as a lack in robustness when applied to different diets as well as the mirroring by cow-individual differences in feed intake (Grelet *et al.*, 2020, Chen *et al.*, 2021b). Although this recent approach might comprise an approximation, the actual heritability of NUE is still unknown.

3.2.2 Increasing NUE of dairy herds on farm

Current on farm strategies to enhance dairy cows' NUE address to the dietary and "non-genetic" animal characteristics, aiming to establish the most appropriate N supply at herd level. Recommendations for farmers include the general reduction of dietary CP, more precise dietary calculations with slightly negative RNBs, as well as the grouping of cows due to their parity (1st and further), lactation stage and milk performance (Martens *et al.*, 2021). Although these recommendations generally increase dairy cows' NUE, the success of these strategies is limited to the N requirement at herd level. The latter is estimated on gross estimates of average feed N intake and

average milk production calculated from bulk tank milk samples (Jonker *et al.*, 2002, Powell *et al.*, 2006, Stamer *et al.*, 2009, Martens *et al.*, 2021).

The dietary N supply at average herd levels leads to the two following hypotheses:

If the NUE variety between individual cows is partly attributed to genetics, individual cows in a herd might be over- or undersupplied by the dietary N provision.

Moreover, a potential predisposition for higher NUE of individual cows remains unexploited by diets, which are estimated for average NUEs.

Selecting individuals with predisposed higher NUE by breeding programs might therefore:

- equalise the NUE at herd level and thus allow more appropriate feeding of all individuals in the herd.
- comprise a management independent tool for an additional increase of NUE on dairy farms.

3.3 Breeding selection of dairy cows

In Germany, the official selection of dairy cows for desirable phenotypes started in the 19th century, when farmers initially documented and subsequently mated purebred animals, and thereby established the pedigree of the contemporary, high yielding Holstein-Friesian breed. To date, the Holstein breed is one of the worldwide leading dairy cow breeds and represents approximately 1.72 million herd book entries and more than 13,000 breeding farms (noted in 2021), the most popular dairy breed in Germany. The Holstein breed is characterised by an average genetic milk performance potential of 10,000 kg milk per lactation (305 days) and 40,000 kg milk per lifetime with a fat content of 4.0% and a protein content of 3.5% (BRS, 2022).

Nowadays, the selection of individuals for specific mating is conducted on the basis of animal specific breeding values. The German breeding value estimation for dairy cows and further livestock is conducted by the independent data centre Vereinigte Informationssysteme Tierhaltung w. V. (vit) and is supervised by the Federal States (Vit, 2022; BRS, 2022; access both: 29.09.2022). Currently, breeding values are estimated for the breeds Holstein, Red Cattle (Angler), Jersey, Red and White dual purpose (RDN) and German Black and White dual purpose (DSN). Breeding values are estimated based on the purpose of a breed specific goal (Vit, 2022; access: 24.09.2022).

In Germany, the potential of an animal in regard of the respective breeding goal is depicted by its RZG. The RZG represents the official total breeding merit index of an individual (BRS, 2022). The RZG is composed of estimated breeding values (EBVs) for individual traits. At present, EBVs for more than 50 individual traits are estimated in the Holstein breed. The EBVs of the individual traits are incorporated

by specific weights into trait complexes, which in turn contribute by relative weights to the RZG. The RZG is reported on the relative breeding value scale with a mean value of 100 and a genetic dispersion of 12 points. The calculation is based on all four- to six-year-old cows, which are listed for the respective breed (Vit, 2022; access: 24.09.2022). The composition of the RZG, hence the trait complexes and their relative weights, is adapted to changing economic and social demands at irregular intervals (BRS, 2022). The current RZG for Holsteins (last update: April 2021) incorporates the trait complexes milk production (36%), longevity (18%), direct health (18%), conformation (15%), fertility (7%), calf vitality and health (3%), as well as maternal: 1.5% and direct: (1.5%) calving traits (Vit, 2022; access: 24.09.2022). Historically, the relative weight of the trait complex milk performance (RZM) continuously decreased in the breeding goal (*e.g.:* percentage of RZG until 1996: RZM= 100% of RZG; 1996: RZM > 50% of RZG; 2008: RZM= 42% of RZG; 2016: RZM= 45% of RZG; 2021: RZM= 36% of RZG) whereas the relative weight of functional traits simultaneously increased (*e.g.:* until 1996: 0% of RZG; 1996: 25% of RZG; 2008: 40% of RZG; 2016: 36% of RZG; 2021: 49% of RZG) (Höltje, 2018). Functional traits represent the entire complex of the so-called non-production traits, which account for the robustness and fertility of an animal (Shook, 2006). In the current RZG, all trait complexes, except for milk production and conformation account for functional traits. The relative weight of the conformation trait complex remained constant at 15% of the RZG since 1996 (Höltje, 2018).

3.3.1 Genomic enhanced breeding selection

The sequencing of the bovine genome (Gibbs *et al.*, 2002, Elsik *et al.*, 2009) and the discovery of DNA markers in the form of single nucleotide polymorphism (SNP) enabled the integration of genomic information on the basis of cost- effective SNP genotype data into classical breeding value estimation systems (Halushka *et al.*, 1999, Meuwissen *et al.*, 2001, Hayes *et al.*, 2009). Like major dairy countries (*i.e.,* USA, Canada, New Zealand, the Netherlands, Nordic countries, France) genomic selection has been implemented routinely in the German breeding system in 2010 (Hayes *et al.*, 2009, Liu *et al.*, 2010, Reinhardt *et al.*, 2011). The classical EBV of a trait, which was based on pedigree, progeny and offspring performances, was thereby extended for a genomic breeding value. The resulting genomic enhanced breeding value (gEBV) is the current, official breeding value for European breeding evaluation systems in dairy cows (Vit, 2022; access: 20.07.22).

For initial gEBV estimation of a single trait, a gEBV is primarily estimated in a reference population. Therefore, the SNP data of the reference population (*i.e.,* genotypes) is associated with the deregressed EBV (*i.e.,* phenotypes) (Seefried *et al.*, 2010, Vit, 2022; access: 20.07.22). The marker effects contributing to the phenotype deviation are estimated in the reference population, summed up for the individual trait and subsequently added up to the gEBV of an animal in correspondence to its individual marker expression (allele frequency). After marker effect estimation in the reference

population, gEBVs for the individual trait can be estimated solely from genotype data in subsequent generations (Hayes *et al.*, 2009).

Until 2019, reference populations were based on bull genotypes and their daughter performance as phenotype data. The strict genomic selection among potential candidates as insemination bulls led to a massive decrease of new daughter-tested bulls for the reference population. Moreover, the bull genotypes in the reference population did not represent the entire genetic range in the Holstein population. That is why German breeding companies initiated the genotyping of entire cow herds, to establish a representative and steadily increasing mixed cow and bull reference population (Reinhardt *et al.*, 2011). Since 2016, genotype data of cows is routinely consolidated and imputed on a standard panel of 45,613 SNPs. Phenotypes for the cow reference population are collected on farm by monthly milk records (milk traits) and by the farmers (functional traits). Since 2019 the reference population for genomic breeding value estimation in Germany is based on a mixed reference population of cows and bulls, which comprises at present (29.09.2022) a number of > 375,000 individuals for the milk trait complex (RZM), > 320,000 for conformation and between 230,000 to 700,000 individuals for functional trait complexes (Table 1; Vit, 2022; access: 29.09.2022).

Table 1. Mixed cow and bull reference population size for gEBV estimation (Vit, 2022).

April 2022[1]	Bulls*	Cows	Total
Milk performance	45.648	329.748	375.396
Conformation	43.28	280.768	324.048
Calving ease	34.426	377.624	412.05
Health (udder)	6.304	232.168	238.472
Calf health	12.043	697.213	709.256

*based on data of > 40 million daughters; [1]access: 29.09.2022

The reference population size expands continuously with each conventional breeding value estimation by the newly added daughter- tested, genotyped bulls and by the cows with phenotype performance. After each conventional breeding value estimation, the marker effects and thus the genomic sum formulas for the gEBVs are re-estimated on the basis of the expanded reference population. The individual cow or bull gEBVs are consequently updated during the main breeding value evaluation releases in April, August and December (Vit, 2022; access: 29.09.2022).

In general, the integration of genomic information increased the accuracy of breeding selection (Hayes *et al.*, 2009). Current gEBVs comprise 69% reliability, compared to classical EBVs (19%) and solely genomic data (67%) (Vit, 2022; access: 21.07.2022). The accuracy of the gEBV is in dependence on the reference population size, the reliability of the EBVs and on the heritability of the trait (Liu *et al.*, 2010).

Although the integration of genomic information massively decreased the number of required phenotypes, the minimum sample size of a mid-, or high-density genotyped reference population for

gEBV estimation of a trait, which is negatively correlated to further traits of the breeding goal, was calculated with 10,000 individuals (Calus and Veerkamp, 2011). Due to the mathematical similarity of efficiency traits to the calculation of energy balance, negative correlations between efficiency traits and energy dependent functional traits (*e.g.:* fertility traits) were evidenced (Pryce *et al.*, 2014, Ferreira Júnior *et al.*, 2018, Spiekers, 2021). Considering an initial gEBV estimation for NUE, a similar reference population size as postulated by Calus and Veerkamp (2011) may therefore be assumed.

3.3.2 Increasing NUE of dairy cows by breeding selection

It was described that cow-individual NUE phenotype data is not available at a large-scale (chapter 2). The difficulty of high-scale phenotype obtainment for breeding value estimation is known from various functional traits (*e.g.:* fertility, fitness). The implementation of these traits into breeding programs was conducted by the utilisation of indicator traits (Shook, 2006). Indicator traits are low-cost measurable phenotypes, thus available at a large scale, and closely related to the goal trait (Berry, 2013). A broad variety of indicator traits and proxies for NUE has been suggested. Due to the non-invasive and routinely obtainment of cow-individual milk samples in the monthly milk record routine on farm, numerous studies focused either on the total milk-N content, the determination of specific milk proteins or on MU as predictors for NUE (Beatson *et al.*, 2019, Lavery and Ferris, 2021). In addition, measurements of N contents or specific N fractions in faeces (Althaus *et al.*, 2013), urine (Kauffman and St-Pierre, 2001), plasma (Kohn *et al.*, 2005, Cantalapiedra-Hijar *et al.*, 2016), hair (Schwertl *et al.*, 2005) and breath (Valente *et al.*, 2018) were approached. Moreover, models to predict cow-individual N intake (Shetty *et al.*, 2017) or NUE from total N in milk (Grelet *et al.*, 2020, Chen *et al.*, 2021b) were published recently.

Similarly, to direct NUE measurements, the determination of specific milk proteins and the approaches on non-milk samples lacked in high-scale phenotype obtainment or by their utility to portray NUE (*i.e.,* N fractions in faeces predicted by near-infrared spectral data (Althaus *et al.*, 2013)). The prediction model approaches might overcome these problems in the future, but up to now the reliability of the models and thus the resulting accuracy to portray NUE was questioned (Chen *et al.*, 2021b). At this time, MU was presumed as the most promising proxy for NUE (Lavery and Ferris, 2021). The routinely obtainment during monthly milk records has led to a huge data availability of cow-individual MU values, which in turn enabled research in the genetic and physiological field.

3.3.3 Milk urea (MU) as an indicator trait for NUE: quantitative genetics

Regarding MU genetics, the literature agrees on moderate heritabilities, ranging in Holstein populations from 0.13 to 0.59 (Wood *et al.*, 2003, Stoop *et al.*, 2007, Stamer *et al.*, 2011, Guarini *et al.*, 2019). MU heritability estimates were similar between 1st and further parities (Jahnel *et al.*, 2021) and

remained constant during the lactation period, which was depicted by high genetic correlations between MU values within one lactation (r= 0.91 in 1st parity cows; r= 0.98 in 2nd parity) (Stamer *et al.*, 2011). The latter aspect enables the utilisation of a repeatability model for variance component and breeding value estimation for MU (Stamer *et al.*, 2011). Although the estimation of genetic correlations between MU and further breeding goal traits in milk differed slightly between studies (Wood *et al.*, 2003, Miglior *et al.*, 2007, Stoop *et al.*, 2007, König *et al.*, 2008, Stamer *et al.*, 2009, Jahnel *et al.*, 2021), there is a general agreement on zero or weak genetic correlations between MU and the yield performance parameters in milk (milk-, protein-, fat-, lactose yield), and weak to moderate positive genetic correlations between MU and the milk percentages of fat, protein and lactose. For example, in the study of (Miglior *et al.*, 2007), genetic correlations between MU and the yield parameters in milk were estimated with r^2= -0.094, r^2= 0.000, r^2= 0.000, and r^2= -0.092 for milk, protein, fat and lactose yield, respectively. The percentages of fat, protein and lactose in milk were genetically correlated to MU with r^2= 0.425, r^2= 0.2000 and r^2= 0.299. The study was based on a random selected dataset with 60,645 test day records of 5,022 Canadian Holstein cows from 91 herds, in parities 1st, 2nd and 3rd and 5 to 305 DIM. These findings suggest only minor consequences for milk performance, if MU would be included in the breeding strategy (Stamer *et al.*, 2011).

In short, the studies on MU genetics revealed that MU fulfils the general requirements for gEBV estimation (Table 2). However, the actual breeding practice does not conduct a molecular-genetic investigation between the breeding goal trait and its potential indicator. Hence, the effects of selecting individuals by indicator phenotypes first become visible in the following generations (post-selection) (Stamer *et al.*, 2011). Conclusively, if MU's genetic profile templates NUE genetics, and thus, selection with MU would select candidates with predisposed enhanced NUE cannot be evaluated by quantitative genetics (*i.e.,* via genetic correlations), due to the data lack in NUE.

Table 2. Requirements for an indicator trait with regard to breeding selection.

Requirements for an indicator trait in regard of gEBV estimation[1]	NUE	MU
Phenotype data availability	?	X
Heritability	?	X
Repeatability (for a breeding value estimation model)	?	X
No undesirable (genetic) correlations to other breeding goal traits	?	X
Phenotypic similarity to the breeding goal trait	X	?
Economic weight	?	?
Phenotypic variety in a population	X	X

[1]referring to (Berry, 2013), X= available and/or true; ?= not available and/or not known yet

3.3.4 MU as an indicator trait for NUE: phenotypic relationship of MU and NUE

The relationship between cow-individual MU and NUE, as well as between MU and the N utilisation and N excretion in urine and faeces was subject to phenotypic study approaches. In a meta-analysis the connection between MU, NUE and urinary-N excretion was evaluated in a dataset of 306 cows. The cows were stalled in 50 experimental farms in Sweden and Finland, repeatedly measured for MU and NUE, and treated with different diets (Nousiainen *et al.*, 2004). Urinary-N phenotypes were calculated by N intake − (milk-N + faecal-N). NUE and urinary-N were subsequently predicted from MU by a regression model. Significantly (*p*- value > 0.001) high predictions were revealed for MU and NUE (r^2= 0.769), for MU and urinary-N yield (g/d) (r^2= 0.735), and for MU yield (g/d) and urinary-N yield (g/d) (r^2= 0.835). Similarly, significant correlations (*p*- value > 0.001) between MU and UU (r^2= 0.63), MU and UU yield (r^2= 0.64), BU and UU (r^2= 0.78) were identified by direct measurements of UU in 24 multiparous lactating cows of the Swedish Red and White dual purpose breed (Gonda and Lindberg, 1994). The relationship between MU, UU and BU was further investigated in dependence on a linear increase of dietary N supply (15% to 21% CP of DM) (Burgos *et al.*, 2007). The dataset comprised twelve multiparous lactating Holsteins in early, mid and late lactation. The results indicated a simultaneous in- and decrease of MU and UU levels as a function of BU until a CP level of 20% in DM was reached. At CP levels > 20% in DM, the UU and MU relationship diverged from linearity and differed between early and late lactation stage. Various models that predicted UU and UU yield from MU were estimated by (Kauffman and St-Pierre, 2001), (Jonker *et al.*, 2002) and (Burgos *et al.*, 2007). These prediction models were applied on a dataset of 133,624 dairy cows (Holstein-Friesians (HF), Jersey (J) and HF×J) to quantify the UU yield reduction by selecting sires in correspondence to MU breeding values (Beatson *et al.*, 2019). Under the assumption that a relationship between MU and UU would remain stable over selection, a reduction of 6.6 kg UU per cow and year was estimated.

This assumption of a stable-over-selection MU:UU ratio was questioned by a study from the Netherlands, where an EBV MU index was established for bulls (Sebek *et al.*, 2007). The evaluation dataset comprised 723 cows with individual repeatedly measured NUE data (entire dataset: 15,720

records) and EBVs for MU that were retrospectively calculated. Cows with low EBV MU had significantly less MU but did not show any differences in NUE or milk performance parameters. It was concluded that breeding for low MU might alter the relationship between MU and urinary-N excretion, since low MU phenotypes were not accompanied by higher N output in milk (*i.e.,* NUE). However, urinary-N was not measured and the authors admitted concerns about the reliability of the EBV estimation, since most of the cows' EBVs were estimated as expectation values only, by the average of the EBV MU of the cow's sire and the cow's dam. In another study, phenotypes of high and low EBV MUs were investigated in 138 cows (Correa-Luna *et al.*, 2021). Significantly, less MU in cows with low EBV MU was not accompanied by less urinary-N (urea + non-urea N fractions) yield (g/d), but by significantly less milk yield. Under low-intensity conditions (milking once a day, 18.6 to 20.4% CP/kg DM) cows with low EBV MU were pictured with lower NUE, due to reduced milk protein yield. These findings contradict the weak or absent phenotypic and genetic correlations between MU and milk performance, which were confirmed in high-scale sample sizes (see above: *i.e.,* Miglior *et al.*, 2007). NUE and UN (g/d) did not derive from direct measurements, but from feed intake estimates and back calculations, which were partly based on total milk-N, which in turn includes MU.

In a current review, it was assumed that selecting based on MU may lead to a different partitioning of N in the body pool (Lavery and Ferris, 2021). Indeed, in a recent study, lactating Holsteins with higher MU values (HMU: 277 ± 9mg/l; n=10) were found to excrete on average significantly (p- value < 0.05) less creatine, than animals with lower MU values (LMU: 189 ± 12 mg/l; n= 10) (Müller *et al.*, 2021). These finding were assumed as a consequence of a significantly lower urinary clearance rate in the kidney of HMU cows. There was no significant difference in UU excretion between HMU and LMU cows in that study.

In brief, the investigations of the phenotypic relationship between cow-individual MU and NUE, N utilisation and N excretion can be summarised by the following facts:

A phenotypic relationship between MU and NUE was significantly demonstrated and most probably due to a linear phenotypic relationship of MU, UU, BU, which remained constant until a specific (very high) CP level was reached (Gonda and Lindberg, 1994, Kauffman and St-Pierre, 2001, Jonker *et al.*, 2002, Nousiainen *et al.*, 2004, Burgos *et al.*, 2007). If this phenotypic relationship was genetically linked, this linearity would have remained constant among selection for low MU candidates and subsequently resulted in a reduction of UU yield (Beatson *et al.*, 2019). The results of (Sebek *et al.*, 2007) and (Correa-Luna *et al.*, 2021) questioned this hypothesis of a stable-over-election MU:UU relationship. Both studies calculated urinary phenotypes (instead of direct measurements) and concerns in the experimental designs were mentioned (EBV MU estimation (Sebek *et al.*, 2007); very high N supply, N intake calculation (Correa-Luna *et al.*, 2021)). Different MU phenotypes were further

observed to differ in renal clearance rates and consequently in the N partitioning of the body N pool (Lavery and Ferris, 2021, Müller *et al.*, 2021).

3.4 Rumen microbiome in dairy cows

The physiological key role of the symbiotic interplay between the cow and its rumen microbiome in the N metabolism of lactating cows was described (chapter 3.1).

The classical knowledge of the rumen microbiology was based on anaerobic, culture-dependent methods (McCann *et al.*, 2014), which were limited to the identification of culturable microorganisms (1% of rumen microbial species (McCabe *et al.*, 2015). The development of DNA- based methods and statistical pipelines for the analysis of –omics data enabled the culture- independent exploration of the rumen microbiome and comprised a milestone in rumen microbial research in the last few years (McCann *et al.*, 2014).

Exemplarily, the 16S rRNA Amplicon sequencing approach, which comprises one of the most popular DNA- based methods in microbiological approaches (Ibal *et al.*, 2022), will be described briefly.

The 16S rRNA Amplicon approach is based on the identification of microorganisms by the amplification of a phylogenetic marker gene (Baker *et al.*, 2003). Most commonly, a small subunit of the 16S rRNA gene is chosen as the marker gene, due to its highly conserved, variable regions (*e.g.*: V3 and V4 region), which comprise phylogenetic sequence identity in microorganisms (McCann *et al.*, 2014). In short, the hypervariable regions of the 16S ribosomal RNA gene, which are present in a microbial sample (*e.g.:* rumen fluid), are amplified by PCR with primers that have sample specific barcodes (Kozich *et al.*, 2013). The 16S amplicons are subsequently sequenced, and the sequence data is subsequently filtered, aggregated into operational taxonomic units (OTU) by sequence identity (most commonly >97% identity) and further aligned to phylogenetic assemblies in microbial databases (*e.g.*: Silva: https://www.arb-silva.de) (Schloss *et al.*, 2009). The 16S rRNA approach enables the taxonomic annotation of the OTU sequences to the genus level and above (*i.e.,* kingdom, phyla, order, family). The microbial information content of this approach is generated by counting the amplified 16S rRNA sequence abundances of the respective taxa (McCann *et al.*, 2014).

3.4.1 Core rumen microbiota

The rumen microbiota constitute bacteria, archaea and fungi, with bacteria accounting for the most abundant group (chapter 3.1, (Von Engelhardt *et al.*, 2015)). Among the bacteria, a core rumen microbiome, which mainly comprised the bacterial genera *Prevotella, Butyrivibrio, Ruminococcus* and the families *Lachnospiraceae, Ruminococcaceae, Bacteroidales* and *Clostridiales* was found across more than 30 different ruminant species and sub-species, including cattle, sheep, and goats

(Henderson *et al.*, 2015). For dairy cows, a rumen microbial core group was identified, which covered between three- and two-thirds of total rumen microbial abundances in a population of more than 1000 cows (Wallace *et al.*, 2019).The core group comprised 454 bacteria, which were assigned to 11 orders and were present in at least 50% of the animals. The results of the study partly overlapped with the results of Henderson *et al.* (2015; see above) and Sasson *et al.* ((2017); n= rumen fluids of 78 lactating Holstein- Friesians).

3.4.2 Rumen microbiota: dietary, host non-genetic, and host genetic effects

Although the major abundance of all microbial core members remained stable across breeds (Holstein Friesian and Nordic Red dairy cattle breed), geographical regions (four European countries) and diets, the abundance of specific core members, as well as less abundant microorganisms, dynamically adapted to environmental changes, which was depicted by shifting of microbial abundances (Henderson *et al.*, 2015, Loor *et al.*, 2016, McCann *et al.*, 2017, Wallace *et al.*, 2019, Cholewińska *et al.*, 2021). For instance, it has been observed, that the host's parity, lactation stage and the host's physiology condition provoke in- and decreases of specific microbial taxa abundances (Cholewińska *et al.*, 2021). Moreover, in similarity to MU and NUE phenotypes, the dietary composition was determined uniformly as the most important factor to influence the rumen microbial composition (Henderson *et al.*, 2015, Loor *et al.*, 2016, McCann *et al.*, 2017).

The study of (Wallace *et al.*, 2019) uncovered a significant correlation between the abundances of 39 ruminal core members with the cows' genotypes (taxonomic assignment to the order, family, or genus level; confidence intervals= 95%). A host genome's effect on microbial taxa abundances was concluded, which was further evidenced by numerous studies of the quantitative genetics field (see below), all of which considered potential for the modulation of the rumen microbiome by future breeding strategies. Due to the high economic and socio-ecological interest, these studies of the quantitative genetics field mainly focused on the rumen microbiome in the context of cow-individual feed efficiency and methane emissions (*e.g.:* Myer *et al.*, 2015, Roehe *et al.*, 2016, Shabat *et al.*, 2016, Difford *et al.*, 2018b, Ramayo-Caldas *et al.*, 2020, Liu *et al.*, 2021, Saborío-Montero *et al.*, 2021, Martínez-Álvaro *et al.*, 2022), but enabled a massive increase in the knowledge on general aspects for host–rumen microbiome –(efficiency) trait axes. The major findings are summarized below.

3.4.3 Rumen microbiota for breeding

The host genome's effects on an animal specific rumen microbial composition was explored by regression analysis between the host's genotype and a dissimilarity matrix of ruminal taxa abundances. Moreover, the host genetic effect on specific taxa abundances was evaluated on the basis of absolute or relative taxa abundance, accordingly. In reference to the principle of heritability estimation, the

host genetic effect was quantified and termed as h^2_B (Difford *et al.*, 2018b, Wallace *et al.*, 2019, Martínez-Álvaro *et al.*, 2022). Based on 16S Amplicon data of 750 rumen fluid samples (lactating Holsteins), filtered for presence/absence in at least >50% of the cows, h^2_B for OTUs (relative abundance) ranged between h^2_B= 0.16-0.44 and between h^2_B= 0.17-0.25 for genera (relative abundance). Heritabilities were significant for 6% of OTUs (in total almost 4000 OTUs) and for 8 out of 144 genera (*p*- value < 0.05; FDR 15%). In another approach, metagenomics sequencing data from rumen fluids of 363 steers (different diets, breeds, seasons, experiments) revealed significant heritabilities for 194 out of 1107 genera with h^2_B= 0.13- 0.61, including 20 genera with h^2_B >0.40 ($P_0 \geq$ 0.95) (Martínez-Álvaro *et al.*, 2022). Considering the large impact of the environment (*e.g.:* dietary composition), which frequently alters the bio-chemical ruminal balance, and consequently the microbial composition (Liu *et al.*, 2021), the host genetic effect on rumen microbial abundances might be interpreted as the host genome's adaptive competence to (re-) establish ruminal housing conditions as a reaction on environmental ruminal imbalances (*e.g.:* by gene-expression) in order to maintain a long-term symbiotic maintenance (Peterson *et al.*, 2007).

It was further postulated that the utilisation of "holobiont" information, which included SNP data as host genome information and 16S rRNA amplicon sequencing or metagenome data as microbial information, enabled a more accurate trait prediction (methane), than the single prediction of the trait by either the host genotype or the rumen microbiome (Pérez-Enciso *et al.*, 2021). This was further evidenced by another study in which the additive genetic effects of host and microbiome without interaction, defined as "holobiability", accounted for 42% to 59% of phenotype (methane) deviation, whereas the respective host genetic and microbial effects explained 15% to 17% (heritability) and 15% to 21% of the trait deviation (Saborío-Montero *et al.*, 2021). This effect of the rumen microbiome on a measured phenotype deviation was estimated on the basis of a dissimilarity matrix of taxa abundances and termed as "microbiability" (Difford *et al.*, 2018a).

3.4.4 Rumen microbiota for MU and NUE

So far, "holobiont" research in the context of NUE or MU phenotypes is rare. It was postulated that about 50% of cow-individual MU variance can be attributed to differences in ruminal fermentation (Hof *et al.*, 1997), whereas the absence of ruminal fermentation differences was suggested to explain the absence of a NUE–MU relationship (NUE determined by N isotopic fractionation, n=16 lactating cows) (Cheng *et al.*, 2014). Moreover, in a recent multiomics approach, based on metagenome sequencing data, the rumen microbiome and the host metabolome were postulated with joint contribution to cow-individual milk protein yields under standardised feeding and stalling conditions (n= 20 lactating cows) (Xue *et al.*, 2020).

With regard to the above findings and the fact that the rumen microbial community was postulated as the major determinant of the N metabolism in ruminants (Bach *et al.*, 2005), it can be assumed that MU and NUE phenotypes are majorly affected by the rumen microbiome, which in term is partly attributed to host genetics.

4 Aim of the studies

The molecular-genetic background of cow-individual N utilisation and N excretion, resulting in cow-individual NUE phenotypes was found to be still unclear. The difficulty of NUE phenotype obtainment established MU as the recently most appropriate indicator trait for NUE. Although MU fulfils the requirements for breeding value estimation its prediction accuracy for NUE has not been finally assessed. Moreover, the rumen microbiome was evidenced as an intermediary key player of cow-individual N metabolisation, that provides partial host genetic dependency.

In order to identify molecular- genetic and ruminal principles of cow-individual NUE and its currently most appropriate indicator MU, four study approaches were conducted which aimed to explore the complex of cow-individual N utilisation and N excretion in its entirety as an important pre-step for modern breeding strategy considerations (Figure 2).

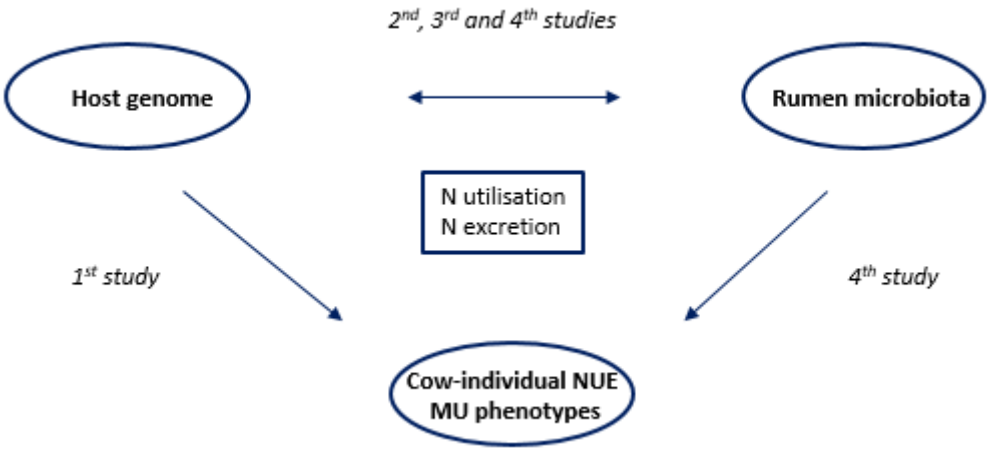

Figure 2. Triangular relationship of host genome, rumen microbiota and cow-individual NUE.

The overall aim of the studies was defined as the identification and assessment of molecular-genetic and ruminal determinants of cow-individual N utilisation and N excretion to enhance the knowledge of cow-individual NUE with focus on MU as an indicator trait for potential breeding strategies.

In specific, the four studies aimed for:

The identification and assessment of genomic regions which are significantly associated with variance in MU and a further eight N excretion metabolites in milk and urine in lactating Holsteins (1st study).

The identification and assessment of ruminal characteristics of cows with predisposed higher compared to lower MU (HMUg *vs.* LMUg) by investigating the rumen microbial community at phylum level and the gene expression in the rumen villi (2nd study).

The identification and assessment of ruminal mechanisms in cows with predisposed higher compared to lower MU (HMUg *vs.* LMUg), fed with either normal or low CP-diet (NP *vs.* LP), respectively, as well as the interaction of host- gene expression and rumen microbial abundances in HMUg *vs.* LMUg cows (3rd study).

The identification and assessment of rumen microbial signatures assigned to a direct genomic breeding value for MU (GBVMU), as well as their relationship to proxies of NUE in milk, urine and faeces in lactating Holsteins (4th study).

5 Cow populations and experimental designs

Data material of the studies was obtained by sampling two different cow populations, which will be titled as “Gut Dummerstorf” (1st and 4th study) and “Blue Cow” (2nd and 3rd study) population in the following chapters. A description of the cow populations and the respective experimental designs is portrayed in Table 3. In brief, for the 1st and 4th study a genotyped (50K MD Genome chip (45,613 SNPs)) population of 371 lactating Holsteins, which covered parities from 1 to 9 and all lactation stages (at least 21 DIM) was sampled for rumen fluid, milk, urine, faeces, and blood on the practice- operating dairy farm Gut Dummerstorf. Sampling was carried out in a maximum time- distance of 24h to the monthly milk record, from which MU values and milk performance parameters were obtained. Samples were taken from each cow once, by separating the cow after morning or evening milking. For 358 cows GBVMUs were estimated by the project partner (Vereinigte Informationssysteme Tierhaltung, vit Verden, Germany).

The 2nd and 3rd study based on a population of 20 non-pregnant lactating Holsteins, which were grouped due to their predisposition for high or low MU values (HMUg *vs*. LMUg) determined by deregressed proofs of an EBV for MU. The cows were grouped in pairs (each pair comprising one HMUg and one LMUg cow). The cow pairs were housed under experimental conditions and fed with either a normal or a low CP diet (NP *vs.* LP), in a 4x5 balanced design (NPxHMUg, NPxLMUg, LPxHMUg, LPxLMUg; each n= 5). After two weeks, the cow pairs were slaughtered. Samples of rumen, kidney, liver, and udder tissue and rumen fluids were obtained directly after slaughter. For the studies, which contribute to this thesis, only the rumen samples (tissue and fluids) were analysed.

Table 3. Cow populations, sampling, sampling designs and sample analyses.

	Gut Dummerstorf	**Blue Cow**
Study approach	1st and 4th	2nd and 3rd
Cow population	371 (genotyped, 50K MD Genome Chip, lactating Holsteins	20 non-pregnant lactating Holsteins
Grouping	GBVMU (4th study only)	Deregressed EBVMU (HMUg *vs.* LMUg) and diet (NP *vs*. LP; 4x5 design, 3rd study only)
Sampling conditions	Practice-operating farm	Experimental farm
Season of sampling	February, March, April 2020	All season (2019 and 2020)
Sampling design	All samples of each cow once, max. 24h to milk record (milk sample)	Sampling directly after slaughter
	Sampling and sample analyses	
Rumen fluid	Rumen microbiota composition via 16S rRNA Amplicon Sequencing	Rumen microbiota composition via 16S rRNA Amplicon Sequencing
Milk	Monthly milk record data of one entire lactation (14 months): urea, protein, fat, lactose* and milk yield; one-time measurements from sampling trial: urea, protein, fat, lactose* and milk yield, total milk-N concentration, short and long chain fatty acids*	Milk urea values of recent five milk records before experimental start; Milk-N metabolites (Müller *et al.*, 2021)
Urine	Concentration of urea, allantoin, uric acid, hippuric acid, creatine, creatinine	Urinary-N metabolites (Müller *et al.*, 2021)
Faeces	N concentration in dry matter	
Blood	Holistic metabolome profile*	Plasma metabolites (Müller *et al.*, 2021)
Rumen tissue		Holistic transcriptome profiling (HiSeq)
Udder tissue		Holistic transcriptome profiling (HiSeq)*
Kidney tissue		Holistic transcriptome profiling (HiSeq)*
Liver tissue		Holistic transcriptome profiling (HiSeq)*

*not utilised data in the studies, which are presented in this thesis

6 Personal contribution to the experimental trials

I declare that my contribution to the experimental trials, the sample obtainment and laboratory analyses was as described below.

"Gut Dummerstorf" sampling trial (1st and 4th study)

- Organisation of the sampling trial "Gut Dummerstorf" and sample analyses (*e.g.:* planning of the sampling procedure, intermediate communication between farm, Institute (three intern research groups), and 3 external project partners (Vereinigte Informationssysteme Tierhaltung w.V. (vit Verden), Qualitätsprüfungs- und Dienstleistungsgesellschaft Mecklenburg-Vorpommern mbH (MQD), RinderAllianz GmbH)
- leading the sampling team (scientists and farm workers)
- sampling all cows (n= 371) for rumen fluid and blood, and partially (supported by the sampling team) for faeces and urine
- preliminary laboratory work for rumen microbiota analysis (DNA extraction, 16S rRNA Amplicon PCR, library quantity and integrity checks), rarefaction of obtained milk record and urinary analyses data (*e.g.:* correction of area under the curves from HPLC urine analyses)

"Blue Cow" sampling trial (2nd and 3rd study)

- sampling of liver, rumen, kidney, udder tissue and rumen fluid post-slaughter
- preliminary laboratory work for rumen microbiota analysis (DNA extraction, 16S rRNA Amplicon PCR, library quantity and integrity checks), and for RNA holistic transcriptome sequencing (RNA extraction, RNA quantity and integrity checks)

7 List of experimental studies and personal contribution

The studies, which contribute to this thesis, are printed in the Appendix. I declare that my contribution to the publications was as described below. A detailed description of my contribution to the experimental trials, which underlie the studies, was given above. The contribution of the co-authors is described in detail in the respective publication.

1st study:

Honerlagen, H., Reyer, H., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., Reinsch, N., and Wimmers, K. (2021). Identification of Genomic Regions Influencing N-Metabolism and N-Excretion in Lactating Holstein-Friesians. *Frontiers in Genetics*, 1241. doi: 10.3389/fgene.2021.69955

Hanne Honerlagen performed, and HR supported the organisation of the sampling trial, cow sampling, statistical analysis, data interpretation and writing of the manuscript. KW supervised the study.

2nd study:

Honerlagen, H., Reyer, H., Segelke, D., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., and Wimmers, K. (2021). Rumen microbiota and host gene expression associate to predisposed milk urea (MU) concentration in Holsteins. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, (WCGALP), Volume Novel Traits, Wageningen Academic Publishers, 2022.

Hanne Honerlagen and HR collected the samples, performed statistical analysis and interpreted the data. Hanne Honerlagen conducted laboratory work and wrote the manuscript. CM, MP, and BK organized and carried out animal husbandry. KW supervised the study.

3rd study:

Honerlagen, H., Reyer, H., Segelke, D., Müller, C.B., Prahl, M.C., Ponsuksili, S., Trakooljul, N., Reinsch, N., Kuhla, B., and Wimmers, K. (2021). Ruminal background of predisposed milk urea (MU) concentration in Holsteins. *Frontiers in Microbiology*, 13. doi: 10.3389/fmicb.2022.939711

Hanne Honerlagen and HR collected the samples, performed statistical analysis and interpreted the data. Hanne Honerlagen conducted laboratory work and wrote the manuscript. HR assisted statistical analyses and interpretation of the data. CM, MP, and BK organized and carried out animal husbandry. KW supervised the study.

4th study:

Honerlagen, H., Reyer, H., Abou-Soliman, I., Segelke, D., Ponsuksili, S., Trakooljul, Reinsch, N., N., Kuhla, B., and Wimmers, K. (2021). Microbial signatures inferred from association with genomic breeding values for milk urea concentration (MU) and their relationship to proxies for N-utilisation efficiency (NUE) in Holstein cows. *Journal of Dairy Science,* Submitted 23rd October 2022.

Hanne Honerlagen performed, and HR supported the organisation of the sampling trial and cow sampling. Hanne Honerlagen performed laboratory work statistical analysis, data interpretation and writing of the manuscript. HR assisted statistical analyses and interpretation of the data. KW supervised the study.

8 General discussion

The implementation of NUE into dairy cows' breeding programs was considered advantageous for socio- ecological, economic and animal's welfare reasons, but is prevented by the lack in large-scale NUE phenotype data. Cow-individual MU values are available at a large scale, which led to considerations on the selection for low MU candidates, aiming to enhance NUE. In addition to MU, proxies of cow-individual NUE are measurable in urine, milk and faeces. The host genome, the rumen microbiome and the interaction between host genome and rumen microbiome were hypothesised to mainly determine cow-individual NUE, its most prominent indicator trait MU and proxies of NUE in milk, urine and faeces of lactating Holsteins. Based on a triangular relationship of host genome, rumen microbiota and N associated phenotypes, which was assumed as holistic, underlying principle of cow-individual N utilisation and N excretion, four study approaches were conducted (Figure 2).

Our studies explored the host genome–trait axis (1ststudy), the host genome–rumen microbiota interaction (2nd and 3rd study), and the triangular host genome–rumen microbiota–trait axis (4th study) (Figure 2). The following chapters provide short summaries of the studies, discussion of the experimental approaches, results and a potential integration of the respective study results. Moreover, future research strategies are formulated and general considerations for breeding strategies on enhanced NUE are contextualised.

8.1 Host genome–trait axis

The initial exploration of the triangular relationship underlying cow-individual NUE focused on the host genome–trait axis.

8.1.1 Short summary (1st study)

In the 1st study the genetic architectures of nine N excretion metabolites in milk and urine were explored by a GWAS approach in 371 lactating Holsteins. For eight N metabolites, significantly associated SNPs were identified. The corresponding genomic regions were discussed in the context of positional (covering the SNP) and functional (biological function of the gene, max. 500kb to the left or right of the SNP) candidate genes, by utilising the gene databases of Ensembl, NCBI, GeneCards and published literature.

8.1.2 Study approach and results

Up to now, our GWAS for urinary phenotypes comprises a unique approach in dairy cows. The collection of cow-individual urinary phenotypes is difficult, time-costly, and generally not feasible in free-stall housed cow herds. Due to the separation strategy of cows directly after milking, our experimental design enabled, in comparison to other studies, the collection of high- scale individual

urine samples (n= 371, *vs. e.g.:* n= 10 cows (Bristow *et al.*, 1992); n= 20 cows (Müller *et al.*, 2021)). Optimal sample size for GWAS approaches cannot be uniformly defined (Korte and Farlow, 2013). It was postulated that some phenotypes were mapped successfully in less than 100 individuals (Atwell *et al.*, 2010). However, for highly polygenic traits low sample sizes are crucial due to the lack in detecting all QTLs underlying a trait (Atwell *et al.*, 2010). The allele frequency and the distribution of (unknown) causal genetic variants in the cow population investigated, as well as the linkage disequilibrium (LD) between the SNPs and the unknown causal variants additionally influence the statistical power (Visscher *et al.*, 2017). The validation of our results in larger populations might uncover additional associated regions.

Considering data accuracy of our study, the imputing of genotypes was conducted in a dataset with more than 800,000 cows, which ensured high quality to the genotypes (Vit, 2022; access: 20.07.22). The phenotypes were obtained “on farm”. Although sampling on a practice-operating farm cannot fully control or quantify environmental factors, the sampling design in our study enabled equal conditions of main factors, such as feed composition, daily routines (milking and feeding), cow stalling, sampling max. 24h to milk record, temperature and climate conditions, which generally distort physiological studies in large populations (Dahlhoff, 2014). However, it was not possible to control or measure the cow-individual feed intake and the urine volume, which were both assumed to bias the phenotypes.

Variability in cow-individual feed intake was considered to influence the amount of bioavailable N (Riemeier, 2004), due to the slightly negative RNB in the feed ration, which in turn would impact the quantity of urea synthesised. High urine volumes were further thought to dilute the urinary metabolites (Cox, 2013).

The linear regression in GWAS approaches is often conducted with adjustment for covariates. Covariate adjustment enables the estimation of the independent effect of a predictor variable, and additionally reduces the residual variability (Wang *et al.*, 2018). The latter improves statistical power (Vansteelandt *et al.*, 2009). However, the adjustment with a heritable covariate is problematic, if its genetic variant was associated with the genetic variant of the target trait (*i.e.,* pleiotropy) (Aschard *et al.*, 2015).

Considering that our GWAS aimed to identify the molecular- genetic architecture of the unbiased, underlying traits, covariate adjustment for milk and urine phenotypes was applied. For milk phenotypes, the cow-individual milk fat percentage was utilised as fixed effect in the models, to account (indirectly) for differences in cow-individual N intake. The cow-individual milk fat to milk protein ratio (FPR) was suggested as a potential indicator of energy balance (Grieve *et al.*, 1986, Heuer *et al.*, 2000, Glatz-Hoppe, 2019). The energy balance in turn is highly genetic correlated to DMI (*e.g.:*

r_g= 0.73; n= 682 lactating Holstein- Friesians in Buttchereit *et al.* (2011)). In regard of a confounding between milk protein percentage and further milk-N phenotypes (Kebreab *et al.*, 2010), not the FPR, but solely the milk fat percentage was utilised as covariate, to adjust (indirectly) for differences in cow-individual DMI variability (*i.e.,* N intake). The milk fat percentage is significantly high genetically correlated to the FPR (r_g= 0.84-0.97 across lactation stages, p< 0.05; (Buttchereit *et al.*, 2011)).

Urine traits were adjusted for covariates by urinary creatinine concentration for differences in urine volumes. Creatinine is excreted independently of dietary intake, endogenous protein metabolism and urinary concentration (Lindberg, 1989, Kraft, 2005), but as a function of muscle mass (Dijkstra *et al.*, 2013). The latter was considered by including the cow-individual body weight as additional fixed effect.

The validation of our results by the biological function of the candidate genes that were identified in our study implied reasonable covariate adjustment. For example, gap junction protein alpha 1 (*GJA1*) was identified as candidate gene for MU. *GJA1* encodes connexin 43, a component of gap junctions that contributes to the diffusion of low molecular weight substances into cells (Graham and Simmons, 2005, Kim *et al.*, 2012, Hu *et al.*, 2013). Gap junctions, which are composed of the connexin 43 component allow the diffusion of molecules <1,200 dalton (Stains *et al.*, 2014). Considering the small molecule size of urea (60.06 dalton) and its diffusion ability across the epithelia alongside concentration gradients (Von Engelhardt *et al.*, 2015, Biorama, 2022), an influence of *GJA1* on the urea pool and subsequently on MU sizes seemed conceivable. Furthermore, the genes *RCAN2, CLIC5, ENPP4, and ENPP5*, were identified as potential candidate genes for UU in our study. All of these genes were attributed a relevant role in the diffusion-driven reabsorption and filtration processes in the nephron (Schwiebert *et al.*, 1994, Jentsch *et al.*, 2002, Vekaria *et al.*, 2006, Wang *et al.*, 2011), which in turn majorly determines UU.

The current literature revealed additional overlaps between some genomic regions identified by our study to similar phenotypes in ruminants. For instance, E74 like ETS transcription factor 2 (*ELF2*) was suggested to explain individual variation in MN in our study, and was further postulated as a major transcription factor, that drives the formation of milk protein in high yielding dairy cows (Maas *et al.*, 1997). A potentially important role of *ELF2* for mammary gland development and milk component variability was further evidenced in a network analysis approach in lactating goats (Zhang *et al.*, 2020). *ELF2* mediates the translation of mRNA to protein and thus might also determine the rate of milk protein synthesis. Milk protein is generally synthesized by the lactocytes, which utilise the influx of AA and dispense unused AA as part of MN in milk. Increased, or decreased milk protein synthesis, induced by *ELF2,* may therefore be hypothesised to impact MN.

Due to the lack of genetic analyses conducted on MN, UU and minor urinary-N fractions, a validation of our results with these (exact) phenotypes by the current literature was not possible. This lack of

genetic approaches on these traits is likely due to the difficult phenotype obtainment, which limits the sample sizes (Shook, 2006, Berry, 2013, Guliński *et al.*, 2016). Standard milk record procedures provide analyses of total milk-N and MU. Total milk-N is composed of protein-N (approximately 95%) and smaller N fractions, the latter constituting urea as the major and most variable N metabolite (approximately 70%) and minor N fractions, which were termed as MN in our study (Rafiq *et al.*, 2016). MN is composed of the same N metabolites that minor urinary-N fractions are (*e.g.:* hippuric acid, creatinine (Bristow *et al.*, 1992)). This might imply a possible shared genetic determination of MN and minor urinary-N fractions. In our study, no overlapping QTLs for MN and minor urinary-N metabolites were identified. However, considering that MN residuals pictured the sum of minor-N fraction variation, instead of the single N metabolites, our results do not principally disprove a shared genomic region. It has been stated that some of the minor urinary-N metabolites comprised relevant potential to influence N_2O emissions by interacting with other N metabolites in the urine composition (Kool *et al.*, 2006). For example, urinary hippuric acid was suggested to massively reduce N_2O emissions by its breakdown products (Dijkstra *et al.*, 2013) and has therefore relevance for further genetic analyses, considering the reduction of N emissions from dairy cows by breeding. A shared genomic region of milk and urinary hippuric acid concentration would enable high-scale phenotype obtainment of cow-individual hippuric acid excretion via milk samples. The determination of hippuric acid in MN and a subsequent GWAS approach is proposed for future research, based on the hypothesis that a shared genomic region between milk and urinary hippuric acid may exist.

8.1.3 Results with focus on MU and future research proposals

In comparison to MN and the urinary traits, MU and MUY are easily and cost-effective available by the monthly milk record data. However, approaches on the genetic architecture of MU are rare, compared to traits, which are incorporated in the current breeding goal (*e.g.:* milk yield, -protein-, fertility traits), and comprise inconsistent results. For example, Bouwman *et al.* (2010) localized four genomic regions on BTA 1, 6, 21 and 23 for one-time measured MU and MUY in Holstein-Friesians by a QTL approach. Three significantly associated SNPs on BTA 1, 7, and 26 were detected for one-time measured MU (no overlap to the QTL on BTA 1 in (Bouwman *et al.*, 2010)) by a candidate gene approach in Brown Swiss cattle (Cecchinato *et al.*, 2014). Another three different genomic regions on BTA 4, 5 and 13 were detected in the same breed and phenotype, but in a GWAS approach (Pegolo *et al.*, 2017). Recently, MU's genomic background was subject to three studies (Ariyarathne *et al.*, 2021, Atashi *et al.*, 2022, Van den Berg *et al.*, 2022). Significantly associated SNPs for repeatedly measured MU were uncovered to differ between parities and lactation stages in > 1,600 Dual-Purpose Belgian Blue cows (Ariyarathne *et al.*, 2021). Moreover, six novel candidate genes were identified in a population of 634 Holstein Friesian, Jersey, and crossbred cows by Single-Step Bayesian Method (Atashi *et al.*, 2022). In addition,

the issue of inconsistent results of MU's genetic architecture was considered by applying two single GWAS approaches on data of two different countries (NZL and Australia) and a subsequent meta-GWAS on data of both countries aiming to identify consistent genomic regions (Van den Berg *et al.*, 2022).

In general, the inconsistency of associated genomic regions suggests MU a complex, polygenic architecture, which might be orchestrated differentially between breeds, lactation stages, and parities. However, differences in the determination of the phenotype (one-time measured *vs.* repeatedly measured MU), the genotype data (different SNP chips, haplotypes) and the statistical approaches (GWAS, QTL-mapping, candidate gene approach, meta- GWAS) might have further prevented consistency of results and impede inter-study validation.

Regarding our results in the context of the evaluating for MU as an indicator trait for cow-individual NUE, our study did not uncover significantly associated SNPs or candidate genes that overlapped between the covariate- adjusted MU model and the covariate- adjusted models for MN and urinary-N excretion metabolites. However, when applying GWAS on the raw models of MU and UU, a shared genomic region on BTA 9 was identified by one significantly associated SNP (additional analysis of MU and UU raw models: SNP rs440003937, BTA 9, 61.34Mb in ARS-UCD1.2; access: 12.08.2022; for MU ($-\log_{10}$ (*p*- value))= 5.91, for UU ($-\log_{10}$ (*p*- value))= 4.18; significance thresholds ($-\log_{10}$ (*p*- value)= 4.31 (suggested) and 5.61 (genome-wide) (Honerlagen *et al.*, 2021)). A number of seven genes is annotated in the respective QTL (SNP position ± 500 kb), which includes Ubiquitin Conjugating Enzyme E2 J1 (*UBE2J1*). *UBE2J1* was targeted in the raw model for UU by an additional significantly associated SNP (rs43608880; ($-\log_{10}$ (*p*- value)= 4.16)). *UBE2J1* has also been identified in a GWAs approach for residual feed intake and daily feed consumption in pigs and was further stated as being involved in the regulation of protein metabolic processes in the same study (Do *et al.*, 2014). In regard of these findings, the identification of *UBE2J1* in the MU and UU raw models may either be interpreted as a shared genomic region of MU and UU phenotypes, which was mirrored by our covariate- adjusted models, for instance due to pleiotropy of the covariate's genetic variant and the MU and UU genetic variants as described above (Aschard *et al.*, 2015). Alternatively, it is possible that the overlapping SNP association in the MU and UU raw models derived solely by the covariate genetic variant (N intake), which confounded both, MU and UU raw models, when covariate- adjustment was not applied (Wang *et al.*, 2018). This latter assumption would support the covariate- adjustment in our study. Pros and cons of covariate-adjustment in GWAS approaches were discussed above: (Vansteelandt *et al.*, 2009, Aschard *et al.*, 2015, Wang *et al.*, 2018) and assumed as plausible for our data. However, in general, validation studies in huge cow populations across different parities, lactation stages, breeds, countries and methods are suggested, considering that the results of our covariate-adjusted models may have mirrored a genetic linkage between MU and UU. Regarding that accurate urine sample obtainment is

limited to sample sizes on experimental farms (Lavery and Ferris, 2021), data sharing might comprise an option for a high-scale UU GWAS approach. However, if feasible sample sizes without major inter-study biases can be achieved, is questionable (Spanghero and Kowalski, 1997).

In regard of more than 300,000 Holsteins that constitute high-scale milk record phenotype data (milk performance and MU) and genotypes, which derived by the same SNP chip that we utilised in our study (see above: reference population for gEBV estimation of milk performance, chapter 3.3.1, Table 1), a practicable, additional exploration of a possibly shared genetic architecture of MU and UU and the further N excretion traits might instead be conducted by the following steps:

- Grouping the reference population for high and low MU phenotypes.
- Allele identification of the SNPs, which we identified as significantly associated to the urine traits in our study.
- Applying Student's t-test on allele frequencies between (extreme) high and low MU groups.

Furthermore, regarding the major wastes of dietary N via faeces, the investigation of faecal-N phenotypes, which were obtained in the "Gut Dummerstorf" cow population may unravel additional genomic regions by a GWAS approach, which could be of particular interest, considering the postulated absences of phenotypic correlations between N excretion in milk and urine to faecal-N. The identification of SNPs that associate with faecal-N might therefore provide additional potential to increase NUE. However, faecal-N is massively influenced by DMI (Stallcup *et al.*, 1975). Hence, the determination of associated genomic regions asks for accurate separation of the faecal-N phenotypes to DMI. Genetic determination of individual faecal-N is therefore suggested if techniques for cow-individual DMI data collection will become available at a large scale.

8.2 Host genome–rumen microbiota axis

The 2nd and 3rd study investigated the host genome–rumen microbiota axis (Figure 2). The rumen microbiota were focused as host genome dependent and interactive key player of the cow-individual N metabolism.

8.2.1 Short summaries (2nd and 3rd study)

In the 2nd and 3rd study the rumen profiles of 20 cows grouped for their predisposition of higher or lower MU (HMUg *vs.* LMUg, n=10 each) were explored by utilising 16S rRNA amplicon sequencing data (microbial data) and holistic transcriptome profiling (host- gene expression) of the rumen epithelia. The 2nd study was conducted as a pre- study, giving a first glance into potential ruminal differences between MU predispositions. The microbial data analysis revealed that HMUg *vs.* LMUg cows shared

the same core microbiome, which was analysed at phylum level (mainly *Bacteroidota* and *Firmicutes*, accounting for 78.7% of total rumen abundances). At family level, however, *Monoglobaceae*, *Ruminococcaceae* and *Acetobacteraceae* were identified with significantly differential abundances between the cow groups. Twenty-eight gene transcript abundances were further identified to significantly distinguish the rumen profiles of HMUg and LMUg cows.

In the 3rd study a further exploration of different ruminal mechanisms underlying MU predispositions revealed that the microbial diversity at OTU level did not significantly differ between HMUg and LMUg cows. A number of 10 significantly differential abundant genera (DAG) was assigned to the previously identified families. The genes with different transcript abundances between HMUg and LMUg rumen tissues were mainly assigned to pathways of the immune system and attributed enhanced immune responses to LMUg compared to HMUg cows. A number of 53 microbial genera and 100 host- gene transcript abundances distinguished the rumen profiles of HMUg and LMUg cows. The analysis of a potential microbe- host interplay revealed 157 significantly correlated microbe- gene pairs and two pronounced microbe–transcript clusters that correlated contrarily. The differential abundant gene transcripts of the respective clusters revealed their contribution to urea cycle (Cluster I) and immune responses (Cluster II). Microbial and host- transcriptome data was further analysed separately for an adaption to low CP diets within the respective cow groups. Whereas some genera were depicted by in- or decreases as consequence of LP feeding in both, HMUg and LMug cows, some genera were identified either in HMUg or LMUg cows, as well as one genus abundance was found to increase in HMUg and decrease in LMUg cows when adapting to LP diet. The analysis of host-gene expression revealed a massive adaption to LP diet in LMUg cows (> 1000 differential expressed genes (DEGs)), that indicated downregulation of immune responses and energy metabolism. HMUg cows showed no significant adaption to LP diets in their transcriptome.

8.2.2 Study approaches and results

The expensive analysis of transcriptome profiling via HiSeq and the fact that cows need to be slaughtered for the obtainment of rumen tissue limit sample sizes for rumen transcriptome profiling. The sample and grouping sizes in our study (n= 20 cows; grouping for MU predisposition n= 10; grouping for MU predisposition x diet: n= 4x5) were in accordance with other publications (*e.g.:* rumen transcriptome of Holstein calves; n= 18; groups: n= 3 (Connor *et al.*, 2013); skin transcriptome of Holstein cows; n= 10, groups: n= 5 (Scholey *et al.*, 2013); blood leucocytes of Holstein cows; n= 16, groups: n= 8; (McLoughlin *et al.*, 2014)). The housing of the cows under standardised conditions (experimental farm) and the valid determination of cow groups (MU predisposition) by deregressed proofs of a breeding value for MU (> 8 million MU records in the estimation model, MU predisposition further confirmed by repeatedly measured MU phenotypes) can be assumed to have reduced

environmental bias, thus increased statistical power and the accuracy of our results. Moreover, potential bias between the groups, for instance due to stress (cortisol level) at sampling time point, was prevented by the experimental design, which comprised the pairwise housing and slaughter of one HMUg and one LMUg cow each, assuming equal occurrence of a potential bias in both groups, thus not biasing the comparative analyses.

Regarding the dietary effect in the context of the short lifespans of mRNA (mRNA turnover) in the tissue (Huth, 2012), the following facts of the experimental design may further be noteworthy:

The dietary passage rate is majorly a function of DMI, which in turn is dependent on the dietary energy concentration (Lebzien *et al.*, 2005). Average energy contents of NP and LP diets fed within the experiment comprised 6.1 ± 0.2 and 6.2 ± 0.2 MJ NEL/kg DM. The average DMI of the cows ranged between 17.0 and 18.1 kg/day (Müller *et al.*, 2021). Based on the calculations in Lebzien *et al.* (2005) a passage rate of 5.2% to 7.7% of DMI per hour can be assumed for the diets provided to HMUg and LMUg cows. The cows were slaughtered and sampled 4h after morning feeding, hence it was assumed that the rumen epithelia was still majorly exposed to the dietary effect at sampling time point, being depictable in the transcriptome.

The differences in rumen profiles identified by our studies were determined by a descriptive approach of differential microbial and transcript abundances. A possible insight into underlying biological mechanisms, which contextualised the differential transcript abundances, was conducted by pathway analysis. Our results depicted LMUg cows with higher immune responses compared to HMUg animals and with downregulated immune response and energy metabolism associated gene cascades, when supplied with LP diet. The predominant role of the immune response in our results, and the evidence for a relationship between immune responses and MU levels (Timonen *et al.*, 2017), immune activity and the microbial composition (Bessman and Sonnenberg, 2016) and immune system and welfare in the context of low CP diets (Sinclair *et al.*, 2014) was discussed within the study.

In the context of immune responses, the predominant role of different gene transcript abundances, which belonged to the bovine Major Histocompatibility Complex (MHC) attracted particular interest in our results. MHC gene transcripts distinguished MU predispositions, were depicted as key players in the host- microbial interaction and were evidenced as potential adapting mechanism to low CP diets (Honerlagen *et al.*, 2022). Differences in the MHC complex between the MU predisposition groups were predominantly driven by the enhanced gene expression of Bovine Leukocyte Antigens (*BOLA*) in LMUg compared to HMUg cows. *BOLA* genes are known as major contributors to the MHC (Ellis and Ballingall, 1999, Lewin *et al.*, 1999). In the context of MU predisposition, it may attract interest, that one specific allele in the *BOLA- DRB3* gene was found to be significantly associated with increased 305-day milk, fat and protein yields in Holsteins (Sharif *et al.*, 1999). Furthermore, genetic variants of the

MHC gene complex were attributed to feed efficiency in Nellore cattle (Ribeiro *et al.*, 2021), and different frequencies and expression levels of MHC genes were associated with feed efficiency in hens (Lamont *et al.*, 1987) and pigs (Gondret *et al.*, 2017).

Transcript abundances of the MHC gene complex were further evidenced as being majorly important for the interaction with the rumen microbiota in our study and were particularly identified to dictate the host T-cell reactivity to biochemical products of commensal bacteria (Bessman and Sonnenberg, 2016). In our study, the MHC genes, *BOLA-B, BOLA-DOB* and *BOLA-DRA* contributed to the correlation cluster II, which distinguished the rumen profiles of LMUg and HMUg cows by positive correlations of their transcript abundances to the respective microbial subset abundances (host-microbiota interaction analysis). The higher abundance of *BOLA-DRA* transcripts in LMUg compared to HMUg cows was strikingly positively correlated to *Roseburia* abundances in the rumen fluid (r= 0.49; *p*- value= 0.03). Interestingly the *Roseburia* genus was mentioned in the context of intestinal health and explored as probiotics for the restoration of beneficial flora (*e.g.:* (Neyrinck *et al.*, 2011, Tamanai-Shacoori *et al.*, 2017, Yang *et al.*, 2022)), whereas the immune response of the host genome is known to maintain the symbiotic relationship with the beneficial flora (Peterson *et al.*, 2007). The *Roseburia* genus was also evidenced to maintain tight junction integrity in the intestine of mice (Tan *et al.*, 2019). In regard of higher abundance of *BOLA-DRA* and *Roseburia* in LMUg compared to HMUg cows, and lower plasma urea pool sizes in LMUg compared to HMUg cows' phenotypes (Müller *et al.*, 2021), it might be speculated if an interplay between *BOLA* gene expression and *Roseburia* abundances promote a commensal microbial community, which lowers the ruminal efflux of NH_3-N across the rumen epithelia by enhanced tissue integrity (Schubert *et al.*, 2021).

The effect of a reduced dietary N-supply on the microbial community and on host specific rumen tissue processes in HMUg and LMUg cows was explored on the hypothesis, that LMUg cows may possess lower N input requirements, thus would be oversupplied by the CP concentration in standard diets. Due to our results, which depicted significant downregulation of immune responses and energy metabolism pathways in LMUg cows, when supplied with LP instead of NP diet, the host's immune activity was considered as a protective role for the host against the infiltration of toxic compounds (*i.e.,* high amounts of NH_3-N) via the rumen wall (Bessman and Sonnenberg, 2016, Tannock and Liu, 2020). With regard to the severe energy requirements of the immune system (Gleeson *et al.*, 2004, Kvidera *et al.*, 2017), the observed downregulation of energy metabolism pathways in LMUg cows when fed the LP ration, is speculated as a result of an altered energy demand to stabilise the RNB. In this context, it is further noteworthy, that HMUg cows showed no adaption in the transcriptome to LP diet. Regarding a potential breeding selection strategy with MU in the future, calculations for the optimal N requirement may ask for a profound validation. This includes but is not limited to accurate N balance studies with a large cohort of LMUg cows and different diets, similarly to the study of

(Martens *et al.*, 2021) which was introduced above (chapter 3.2.1.1; > 11,000 Holstein cows housed in an experimental farm, cow-individual NUE measurements, exploration of a gradual dietary RNB decrease.

Among the pronounced identification of immune system pathways, which discriminated the rumen profiles, gene cluster I in the host-microbiota interaction analysis revealed significant enrichment of the urea cycle due to higher expression of Arginase 1 (*ARG1*) in LMUg compared to HMUg cows. Arginase is an enzyme that catalyses the last reaction step in the urea cycle by facilitating the synthesis of arginine into ornithine and urea. This result was discussed within the study, implying that LMUg cow's rumen tissue might facilitate the breakdown of arginine into ornithine and urea. In general, the urea cycle is predominantly conducted by hepatocytes in the liver (Emmanuel, 1980, Razmi *et al.*, 2005). However, (Beck *et al.*, 2009) stated the expression of *ARG1* also in the rumen epithelium of cattle, which implies catabolism of N-metabolites via the urea cycle in the rumen wall. Evidence of considerable arginase activity was further postulated by Emmanuel (1980) and by Aminlari and Vaseghi (1992) in the rumen epithelium, and was also observed in the kidney of domestic species (Aminlari and Vaseghi, 1992), and in the mammary gland of lactating cows (Shaw and Petersen, 1938). In regard of these findings analyses of the transcriptome profiles of kidney, liver and udder tissue, which were obtained for LMUg and HMUg cows within the project, may deserve a specific focus on *ARG1* transcripts and on a potential enrichment of the urea cycle by comparing LMUg to HMUg cows in ongoing research. Moreover, considering a potential difference in the N partitioning between cows with LMU and HMU phenotypes, which was hypothesised by Müller *et al.* (2021), the transcriptome analyses of the different tissues (kidney, udder, liver, rumen) and a subsequent across-tissue data integration may uncover underlying principles of potential differences in the flown partitioning between HMUg and LMUg cows.

Regarding the interpretation of results from pathway analyses, it should generally be considered that significant pathway enrichment is in dependence on the number of genes loaded, and on the number of genes annotated for the respective pathway. Although, optimal sizes of gene sets for enrichment analysis have not been determined yet (Holmans, 2010), the enrolled gene list of cluster I in our study, which depicted enhanced urea cycle due to *ARG1* transcript abundance, comprised 24 annotated genes, which is small compared to the other gene lists enrolled by our study (68 to 1220 genes) or by other studies (*e.g.:* 500 genes (Scholey *et al.*, 2013), 3250 genes (McLoughlin *et al.*, 2014)). Moreover, although some bovine databases are available, these comprise only a few more than 3000 functionally annotated genes (Fortes *et al.*, 2010) (*e.g.:* compared to > 50,000 annotated pathways in Ingenuity Systems (IPA), http://www.ingenuity.com), making them not feasible for enrichment analyses. That is why pathway analyses for livestock are – to date – conducted on the basis of human and mouse assemblies (Curwen *et al.*, 2004, Notebaart *et al.*, 2006), although it was evidenced that species specific

gene annotations may limit their detection and annotation in the underlying assemblies (Karisa *et al.*, 2013). Our results of the pathway analyses and specifically the significant enrichment of the urea cycle may therefore be interpreted as a primary understanding of differential rumen profiling between MU predispositions, diets and host-microbe interactions, which asks for future exploration if underlying tools become more applicable for species-specific data analyses.

8.2.3 Over study discussion: integration of genomic regions and and transcriptome profiling for future research

From our results, particularly *ARG1* and *BOLA* gene transcripts were emphasised as potential molecular determinants of cow-individual N utilisation and N excretion. Hypothesising that genetic variants (*i.e.,* SNPs) may underlie their differential transcript abundance in HMUg compared to LMUg cows, a further exploration of these genes by a targeted sequencing approach in the Gut Dummerstorf cow population is proposed. DNA material of the population is available and sequencing companies offer various methods for targeted sequencing approaches as a standard procedure (*e.g.:* amplicon sequencing, target enrichment, targeted gene panels; Eurofins Genomics (https://www.eurofinsgenomics.eu); Illumina (https://emea.illumina.com/); access both: 11.10.2022)).

In principle, a targeted sequencing approach isolates the region of interest from the entire genome in a DNA sample, for instance by hybridization to a targeted probe. The target genomic region is subsequently sequenced. Compared to whole genome sequencing, this approach enables accurate and low-cost sequencing of specific genomic regions. The sequencing of the targeted regions can either be conducted by next-generation sequencing (NGS), which is cost- efficient, if an entire gene panel of interest exists, or alternatively by Sanger sequencing, which is considered as a cost- efficient and accurate sequencing method, if the sum of sequenced base pairs is low (Eurofins Genomics (https://www.eurofinsgenomics.eu); Illumina (https://emea.illumina.com/); access both: 11.10.2022)). Considering our results, NGS might be advantageous if additionally, to *ARG1* and *BOLA* genes, a set of DEGs is extracted from our results. Sanger sequencing is suggested, if *ARG1* and *BOLA* genes are small genes and no additional candidate genes are investigated.

The aim of a targeted sequencing approach with *ARG1* and *BOLA* genes would be a potential identification of SNPs in *ARG1* and *BOLA* genes, which associate with phenotypes of N utilisation and N excretion. Therefore, a comparison of *ARG1* and *BOLA* gene sequences between the Gut Dummerstorf cows grouped for high and low N excretion metabolites in urine and milk (traits of the GWAS study), faecal-N or genomic breeding values for MU (GBVMU, 4th study).

In addition to targeted sequencing, SNP discovery by RNA-Seq data was postulated as an efficient and cost-effective approach, which was conducted successfully for transcribed genomic regions in livestock

species (Jehl *et al.*, 2021). This approach could be applied on our data by extracting the transcripts of the *ARG1* and *BOLA* genes from the holistic transcriptome profiles in the Blue Cow population. The transcript sequences would be investigated for differences (*i.e.,* SNPs) between HMUg and LMUg cows. In general, SNP discovery by RNA-Seq is limited to the discovery of SNPs, which are localised in transcribed regions (Cánovas *et al.*, 2010), hence, SNPs in introns, SNPs in promoter regions and in general all genetic variants that are not localised in the transcribed region, but influence the transcript abundance, cannot be identified. Moreover, allele-specific imbalance in the transcriptome profile (*e.g.:* due to silencing of maternal or paternal allele) can bias SNP discovery by RNA-Seq approaches (Stevenson *et al.*, 2013). Despite these facts, the availability of the transcriptome profiles of the rumen tissue, as well as from udder, kidney and liver tissues (analysis in the future) enables this approach being conducted without any extra (sequencing) costs. Thus, SNP discovery by RNA-Seq is considered as an efficient approach for the future exploration of *ARG1* and *BOLA* genes as potential candidate genes for cow-individual N utilisation and N excretion.

The integration of gene expression data and GWAS studies was proven to identify causal contributions of candidate genes to diseases in humans (*e.g.:* (Bell *et al.*, 2010, Musunuru *et al.*, 2010, Small *et al.*, 2011)). A data integration approach of the QTLs identified by our GWAS and the DEG dataset identified by the 2nd and 3rd study, may provide a deeper insight into causalities of the host genome driving the cow-individual N metabolism. Potential overlapping results, which might be uncovered by a profound integration approach, may be speculated for the QTLs that were suggested for MU and MUY by the GWAS and the DEG list, considering that the grouping factor underlying the DEGs was the MU predisposition. However, the hypothesis of a different N partitioning, as well as significantly higher urinary creatinine concentration in cows with LMU compared to HMU phenotype (Müller *et al.*, 2021), implies that the transcriptome profiles of LMUg and HMUg cows could also comprise DEGs, which are relevant for other N excretion metabolites than MU, thus data integration is proposed for all QTLSs(*i.e.,* all N metabolism related traits) that were identified in the GWAS.

A first step for the data integration of our studies may comprise the genomic annotation of the DEGs identified in the 2nd and 3rd study and the subsequent exploration of potential overlaps to the genomic regions, which were identified as QTLs for N excretion metabolites in milk and urine by the GWAS. Similarly, the annotation of results from the transcriptome profiles in kidney, udder and liver tissue is suggested for comparison to the QTLs in the future.

So far, potential overlapping results of the 1st to the 2nd and 3rd study were broadly explored by comparing the suggested candidate genes identified by the GWAS to the gene identifiers of the DEGs, which differentiated HMUg and LMUg cows in the 2nd and 3rd study. This broad screening did not uncover any overlaps. However, this can be due to various aspects, such as differences in the

genotypes between the two cow populations, the mirroring of candidate gene transcripts due to the holistic approach (noise) and the complexity of the trait (*e.g.:* a number of genetic variants, with each having a small effect on a trait), tissue specific expression *etc.* (Huang, 2015, Uffelmann *et al.*, 2021). Moreover, considering that the majority of SNPs, which has been identified for complex traits map to non-coding intergenic and intronic regulatory regions (Hindorff *et al.,* 2009; Maurano *et al*., 2012], it is conceivable that regulatory candidates, which were suggested by our GWAS (*e.g.:* E74 like ETS transcription factor 2 (*ELF2*)), may contribute to the measured trait deviation without abundance in transcriptome profiles. Similarly, also a protein- coding candidate gene may have not been depicted the transcriptome profiles, due to a short and time point specific transcript abundances within a gene cascade (Huang, 2015). For example, the QTL on BTA 7, which was identified by our GWAS, comprised several genes with functional annotations in extracellular matrix formation and degradation (*Col23A1, B4GALT7, PHYKPL*), in nuclear ribosomal processes (*HNRNPAB, NHP2, RMND5B*) and in intracellular transports to the Golgi complex (*TMED9*), which may eventually imply time-point specific and subsequent transcript abundance being not present at sampling time point.

8.3 Host genome–rumen microbiota–trait axis

Although the validation of microbial study results by inter- study comparison might be partly impeded due to environmental biases (*e.g.:* feed ration, stalling factors), the underlying data (*e.g.:* 16S *vs*. metagenome *vs.* cell culture), different data filtering procedures and the still rare knowledge of rumen microbial processes, the microbial genera identified in the 2nd and 3rd study partly overlapped with other findings in N efficiency phenotypes of ruminants (goats, yaks, steers, beef). Some microbes attracted further interest as part of the microbial signature inferred from GBVMU selection, which was identified by the 4th study. The 4th study explored the entire triangular relationship that had been initially established to determine cow-individual N utilisation and N excretion (Figure 2).

8.3.1 Short summary (4th study)

In the 4th study, the Gut Dummerstorf cow population was grouped for the upper and lower 15% of an estimated genomic breeding value for MU (GBV_{HMU} n= 29, GBV_{LMU} n= 30; GBV_{MED} n= 299). MU phenotypes differed by 66 mg/l MU ($p < 0.001$) between the extreme cow groups (GBV_{HMU} *vs.* GBV_{LMU}) MU and UU correlated significantly moderate ($r= 0.25$; $p < 0.05$), in the entire cow population, faecal-N did not. The microbial analyses in the 4th study were conducted in a two step- approach by focusing the rumen microbiota (i) as dependent attribute of the host genome (host genome–rumen microbiota axis), and (ii) as potential drivers of NUE associated phenotypes (rumen microbiota–trait axis). The microbial data was investigated in a diversity approach at OTU level (Simpson index) and for differential abundances at genus level between the extreme GBVMU groups. No differences between the groups were revealed at OTU level, but 24 microbial genera (partly) distinguished the rumen

profiles of the extreme GBVMU groups and were consequently suggested as a microbial signature inferred from genomic breeding selection for MU. The microbial signature comprised significantly higher abundances of the ureolytic genus *Succinivibrionaceae UCG-002* in GBV_{LMU} cows. GBV_{HMU} animals hosted higher abundances of the genera *Clostridia unclassified* and *Desulfovibrio*, both of which are prominent for their hyper-ammonia producing species (HAB bacteria). These findings were in accordance with the results of the 2nd and 3rd study in the "Blue Cow" population (higher abundance of *Succinivibrionaceae UCG-002* in LMUg cows; higher abundance of HAB bacteria in HMUg). Moreover, an additional four genera of the microbial signature in the 4th study were also identified to distinguish the rumen profiles of HMUg and LMUg cows in the 2nd and 3rd study.

The respective taxa abundances of the microbial signature were further investigated for their relationship with proxies of NUE in milk, urine and faeces in the entire cow population by correlation analysis (n= 358 cows). Three genera of the *Lachnospiraceae* family revealed significant correlations to MU values and were emphasised as considerable players in the host genome–rumen microbiota–MU axis. The significant correlations of *Prevotellaceae UCG-003*, *Anaerovibrio*, *Blautia* and *Butyrivibrio* abundances with MU measurements, MN and the N content in faeces further suggested their contribution to genetically determined N utilisation in Holstein cows.

8.3.2 Host genome–rumen microbiota axis: results and future research proposals

In general, the modulation of specific microbial abundances is thought to improve phenotypic performance of dairy cows (Weimer *et al.*, 2010, Martínez-Álvaro *et al.*, 2022). Indeed, the manual transfer of rumen fluid from bisons successfully altered the rumen microbial composition in beef heifers and was accompanied by an enhanced protein digestibility and N retention in the heifers (Ribeiro *et al.*, 2017). However, in addition to the fact that manual rumen fluid transfer is not practical at operating dairy farms, the study of Weimer *et al.* (2010) revealed, that lactating Holsteins re-established a host-specific rumen microbiome after near-total exchanges of ruminal contents. These findings were assumed as the host genome's influence to maintain an individual microbial composition. The host genome's effect on the rumen microbiota was further evidenced by considerably moderate to high heritability estimates for entire microbial communities and for specific taxa abundances (Roehe *et al.*, 2016, Sasson *et al.*, 2017, Difford *et al.*, 2018b, Wallace *et al.*, 2019, Martínez-Álvaro *et al.*, 2022). Moreover, the different abundance of the microbial signature (24 genera), which was identified by our 4th study as being inferred from GBVMU selection, is suggested as host genome dependent.

In regard of these findings, the modulation of the microbial composition by targeting genomic regions on the host genome, which associate with beneficial microbiota abundances, might comprise a considerable pre-step to influence phenotypes by future breeding selection strategies with microbial

modulation. Due to the major role of the rumen microbiome for the N metabolism, such a "microbial GWAS" approach is considered particularly important for a breeding strategy on enhanced NUE.

A microbial GWAS approach has been recently conducted in 1,028 pigs (Bergamaschi *et al.*, 2020) and in 40 dairy heifers (Golder *et al.*, 2018). Whereas microbial phenotypes in pigs can be easily obtained by faecal samples, the effort to obtain rumen fluids obviously impedes large-scale sample sizes and thus statistical power of microbial GWAS in ruminants. Considering this aspect, the sample size of our data (n=371 Holsteins) might enable comparably high statistical power for a bovine microbial GWAS. Moreover, the identification of the microbial signature inferred from GBVMU selection and the rumen microbial profile which distinguished HMUg and LMUg cows in the 2nd and 3rd study enables a focused approach of specific microbiota with regard to the cow-individual N metabolism.

For a microbial GWAS approach in the Gut Dummerstorf cow population with focus on MU and further NUE associated phenotypes the following steps are proposed:

- Establishing the microbial phenotypes, by extracting the count data of the microbial families and genera identified in studies 2-4 in the Gut Dummerstorf microbial dataset.
- Normalisation of the count data either by variance-stabilizing transformation (vst) or by relative abundances in the rumen fluid. Variance-stabilizing transformation provides a normalisation of count data in dependence on the library size, but independent to the rest of taxa abundances in a sample (Anders and Huber, 2010, Lloréns-Rico *et al.*, 2021). The relative abundance comprise a normalisation of the respective taxa counts in the context of the entire microbial composition in each sample (Turpin *et al.*, 2016, Weiss *et al.*, 2017). Microbial GWAS based on vsts or on relative abundances might therefore represent two individual, informative approaches and are both suggested for application.

If significant microbe-SNP associations can be identified a further exploration of the respective genomic regions may include:

- Their investigation in the bovine assembly for functional or positional candidate genes and in the published literature.
- Their comparison to genomic regions identified by the 1st study.
- A particular interest for genomic regions, which associate with microbial abundances that in turn were considerably related to faecal-N concentration (correlation analysis between microbial signature and proxies for NUE, 4th study). Considering that proxies for NUE in milk and urine were the consequence of host-assimilated N, whereas faecal-N majorly derives by non-digested N (Stallcup *et al.*, 1975), genomic regions which associate with faecal-N related

microbes may reveal additional informative content on predisposed ruminal-intestinal mechanisms of cow-individual NUE.

- The comparison to ruminal gene- expression patterns (DEG and cluster I and II) that were attributed to MU predisposition (2nd and 3rd study).
- The investigation of allele frequencies of associated SNPs in the extreme GBVMU groups (GBV_{HMU}, GBV_{LMU}).

SNPs identified by a microbial GWAS might be considered as potential targets of cow-individual N utilisation and N excretion and considered for breeding strategies in the future. However, it should be noted that an association study does not provide proofs of causality. Microbial GWAS results for weak or not significantly heritable genera or OTUs may therefore be questioned. The findings of (Difford *et al.*, 2018b) indicated significant h^2_B for 6% out of (almost) 4,000 OTUs and for 8 out of 144 genera (*p*-value < 0.05; FDR 15%). Moreover, wide h^2_B ranges between single OTUs and genera were evidenced (h^2_B= 0.16-0.44 OTU; h^2_B= 0.17-0.25 for genera (relative abundance) (Difford *et al.*, 2018b). Regarding the microbial families and genera identified by our studies, for some of them, significantly moderate to high h^2_B were postulated (*e.g.*: *Acidaminococcus* h^2_B= 0.54 (Martínez-Álvaro *et al.*, 2022), *Prevotella* and *Butyrivibrio* h^2_B= 0.23 to 0.33; *uncultured Prevotellaceae* h^2_B= 0.32 to 0.45 (Wallace *et al.*, 2019, Martínez-Álvaro *et al.*, 2022), $P_0 \geq 0.95$). Their considerable h^2_B as well as the fact that these genera were identified in both of our cow populations (*i.e.,* distinguished MU predispositions (3rd study) and inferred from GBVMU selection (4th study)) a host genome's effect and a potential influence on MU and further proxies for NUE seems conceivable. Thus, these genera are particularly suggested for a microbial GWAS approach.

Due to the difficulty of rumen fluid sampling and the resulting lack in high- scale rumen microbial phenotypes the integration of microbial phenotypes into breeding programs is not possible yet. Sample sizes for initial gEBV estimation were calculated with at least 10,000 individuals (Calus and Veerkamp, 2011). A study with comparable sample size for microbial phenotypes has not been published yet (*e.g.:* n= 358 cows in our study; n= 363 in (Martínez-Álvaro *et al.*, 2022); n= 750 in (Difford *et al.*, 2018b); n= 1000 in (Wallace *et al.*, 2019)). However, adding microbial information was proven to enhance the model prediction accuracy for methane phenotypes in dairy cows (Difford *et al.*, 2018b, Pérez-Enciso *et al.*, 2021). In regard of this fact as well as in consideration of the massive increase in microbial research in the last few years, a soon implementation of microbial information into breeding models may not be far-fetched. In this context, microbial GWAS could contribute as an initial approach to future "holobiont" breeding strategies.

8.3.3 Rumen microbiota–trait axis: results and future research proposals

Whereas it is possible to quantify the host genome's effect on (i) the entire microbial composition by h^2_B estimates for a microbial dissimilarity matrix, and on (ii) specific microbial abundances by h^2_B estimates for single OTU abundances at different taxa levels, the rumen microbial effect on a specific trait (microbiability) is generally estimated by regression analysis between a microbial dissimilarity matrix, which is based on total rumen microbial abundances and the measured trait (Difford *et al.*, 2018a). The effect of specific taxa abundances on the trait cannot be estimated.

In regard of the fact, that N is required by all rumen microbes for growing processes (Von Engelhardt *et al.*, 2015), it seems conceivable that the majority of rumen microbes contributes to N associated phenotypes. Thus, microbiability estimation might enable a reasonable approximation to what extent the rumen microbiota contributes to cow-individual MU and further proxies of NUE.

However, at this time point a dissimilarity matrix and thus the results of microbiability estimation, is a descriptive result of count data. Various normalisation methods were approached, which partly overcome statistical issues of microbial count data, such as the comparability between OTU variance of very high compared to very low abundant taxa (Love *et al.*, 2014, Pérez-Enciso *et al.*, 2021). Nonetheless, yet it is not possible to account for the whole biological complexity of microbial abundances driving a trait (Liu *et al.*, 2021).

Our study results suggest that some genera are particularly important for cow-individual MU and proxies of NUE. In regard of a microbiability estimation based on a dissimilarity matrix, the influence of these potentially major players may be underestimated considering their differential abundance as being only a small part of total ruminal abundance variance. A recent review suggested holistic approaches, which integrate host genome, host transcriptome and host metabolome in addition to the metagenome, metatranscriptome and the microbial produced metabolome and the host metabolome for future approaches to enhance knowledge on the complexity of "holobionts" (Liu *et al.*, 2021). However, up to now the high costs of these -omics approaches hamper the number of studies and high-scale sample sizes. A currently published multiomics approach was conducted on metagenomics sequencing data and revealed that cow-individual milk protein yield was influenced by the metagenome, the annotated metagenomics function, as well as by the microbial and host's metabolome. The study was conducted on samples of the 20 most extreme (very high, very low, n= 10 each) phenotypes from a population of 374 cows (Xue *et al.*, 2020). Furthermore, in a dataset of 48 beef cows, metagenomics analyses identified a breed effect on the microbial composition, which did not affect the specific microbial functions, the latter being determined by the metatranscriptome data (Li *et al.*, 2019). Contrarily, Martínez-Álvaro *et al.* (2022) stated considerable heritabilities for rumen microbial functions, which in turn would imply that breed effects should not be excluded as affecting

microbial performance. These few publications might give an initial glance into the complexity of "holobionts", but also into the difficulty of valid results, due to low sample sizes and the data complexity.

Considering our results, the overlapping findings of some specific microbes in both cow populations, the identification by two different statistical methods within one cow population, the biological assessment by the published literature, and the correlation analyses with MU and NUE proxies may have provided a comparably valid identification of specific microbial key players in the host genome–rumen microbiota–trait axis. Exemplarily, the identification of *Cholstridia unclassified* as one of these ruminal key players is elucidated below:

- *Cholstridia unclassified* revealed considerable mean relative abundances in the Blue Cow population by accounting for 0.2% of total ruminal abundances (790 genera) in all cows (n= 20) and in the Gut Dummerstorf population, covering 0.3% of total ruminal abundances (1,064 genera) in all cows (n= 358).

- *Cholstridia unclassified* was postulated as a genus, to which various hyper-ammonia producing (HAB) species were assigned to (Paster *et al.*, 1993, Bento *et al.*, 2015, Libera *et al.*, 2021). As a result of their high proteolytic activity, HAB bacteria cause steep increases of ruminal NH_3 levels that increase the NH_3 efflux into the bloodstream (Hartinger *et al.*, 2018), which in turn was assumed to enhance the necessity for N detoxification into urea by the host, and would consequently increase MU (Honerlagen *et al.*, 2022). HAB bacteria were indeed stated to negatively influence the NUE of ruminants (Hartinger *et al.*, 2018), which was evidenced by their higher ruminal abundances in Nellore steers with low NUE phenotypes (Bento *et al.*, 2015) and in low N recycling efficiency phenotypes of beef cattle (Alves *et al.*, 2020).

- Our results identified *Cholstridia unclassified* as more abundant in genetically determined HMU hosts in both cow populations (*i.e,* HMUg cows; GBV_{HMU} cows).

- *Cholstridia unclassified* was identified as differential abundant by one statistical analysis (sPLS-DA) in the Blue Cow population and by both statistical analyses (DESeq2 and sPLS-DA) in the Gut Dummerstorf population. In the correlation analyses between microbes and rumen transcriptome of HMUg and LMUg cows (Blue Cow population), *Cholstridia unclassified* was assigned to Microbial Group A. The abundances of microbial Group A were positively correlated to gene cluster I (pathway enrichment: urea cycle) and negatively correlated to cluster II (pathway enrichment: immune responses).

- *Cholstridia unclassified* abundances were further depicted with significant (adj. *p*- value < 0.05) positive correlations to MU (r= 0.18) and with significant negative correlations to total milk-N

(MN: r= -0.17), milk protein yield (MPY: r= -0.16) and faecal-N concentration (FaecN r= -0.16) in the Gut Dummerstorf cow population (n= 358).

Taking the results of our study approaches together, high ruminal abundances of *Chlostridia unclassified* may be determined by host genetic predisposition for higher MU. This may lead to enhanced ruminal NH_3 efflux, which contributes to the host assimilated N pool that cannot be utilised for milk protein build up (negative correlation to MN and MPY) but needs to be detoxified to urea (positive correlation to MU; higher abundance in HMUg and GBV_{HMU} cows). Enhanced ruminal NH_3 host absorption would generally decrease the faecal-N losses (negative correlation to FaecN). Our studies promote *Chlostridia unclassified* as a genus whose ruminal abundance is (partly) under host genetic control and whose higher abundances may contribute to cow-individual higher MU and eventually lower NUE phenotypes.

8.4 Applied breeding selection: difficulty of efficiency traits

Due to the fact that they are ratios of input and output, nutrient efficiency traits depend on the cow-individual DMI, which in turn majorly affects the energy balance of high yielding cows (Buttchereit *et al.*, 2011, Spiekers, 2021). Per definition of the formula, low DMI is related to high efficiency. Hence, "black box" breeding on high efficiency may result in selection for a trait that is like selecting for reduced (or negative) energy balance (Pryce *et al.*, 2014, Spiekers, 2021).

This conflict between energy and efficiency, which derives by the similarity of the formula for efficiency and DMI, may also be assumed for NUE in regard of the following aspects:

In a standard diet, which contains 15% CP, the default N to feed protein conversion factor of 6.25 (Jones factor) and the default N to milk protein conversion factor of 6.38 (applied for milk record data), the dependency between NUE (N in milk (g)/N intake (g)) and DMI can be formulated as:

$$NUE = \frac{\mathrm{MY} \times 1000 \times (\mathrm{MP} \div 6.38)\ \mathrm{g\ N}}{\mathrm{DMI} \times 1000 \times (0.15\ \mathrm{CP} \div 6.25)\ \mathrm{g\ N}}$$

[1]MY= milk yield in kg; [2]MP= milk protein in %CP of 1kg milk; [3]DMI= dry matter intake in kg; [4]0.15 CP due to 15% CP in 1kg DM

The formula implies that a DMI increase leads to a decrease of NUE as a function of non-absorbed N (which is desired). Certainly, NUE additionally decreases when DMI increases due to the feed to milk conversion factor of < 1 (VandeHaar *et al.*, 2016). Assuming exemplarily the feed to milk conversion ratio of 1.95kg feed per 1kg milk, which was recently estimated for lactating Holsteins (Krpálková *et al.*, 2021), NUE can be formulated as:

$$NUE = \frac{\mathbf{1} \times \mathrm{MY} \times 1000 \times (\mathrm{MP} \div 6.38)\ \mathrm{g\ N}}{\mathbf{1.95} \times \mathrm{DMI} \times 1000 \times (0.15\ \mathrm{CP} \div 6.25)\ \mathrm{g\ N}}$$

[1]MY= milk yield in kg; [2]MP= milk protein in % CP of 1kg milk; [3]DMI= dry matter intake in kg; [4]0.15 CP due to 15% CP in 1kg DM; [5]feed to milk conversion ratio: 1.95kg feed per 1kg milk (Krpálková *et al.*, 2021)

Considering the high average DMIs of 23kg DM/d in Holsteins (Martens *et al.*, 2021), this multiplication factor in the denominator massively decreases NUE of high yielding cows, due to feed-to-milk conversion < 1, but independent to the N metabolism of the cow. Moreover, daily measurements of cow-individual DMI and MY were analysed in a data set of >2.200 dairy cows (Holsteins (75%), Simmental (20%), Brown cattle (5%)), which were stalled in ten experimental farms (Gruber *et al.*, 2006). Repeatedly, daily measured DMIs (28 weeks) ranged from 15.9 kg DMI/d to 20.8 kg DMI/d (86% to 111% of average DMI (18.7 kg/d) between the cows. Daily DMIs of individuals ranged from 5 kg DMI/d to 32 kg DMI/d. Increases and decreases of DMI were accompanied by increases or decreases of MY, which were however not linear. In regard of these findings, it might be considered if inter-individual NUE variation that was observed by applying the NUE formula may rather represent inter-individual variation in the feed to milk conversion (*i.e.,* feed efficiency) than inter- individual NUE variation.

In addition to these issues of the NUE formula, which mainly derive by the definition of the denominator (DMI similarity), the definition of the NUE formula's numerator "N in milk" may further be discussed. "N in milk" includes protein-N as major part (MP 95%), but also the milk's NPN (MU and minor N fractions, 5%). Although cow-individual variability in MU and minor N fractions might quantitatively not affect total N in milk, the potential of the cow to utilise or excrete assimilated N (protein-N *vs.* urea-N and minor N fractions) is not displayed by the NUE formula's numerator.

8.4.1 Novel goal trait: the "N efficient holobiont"

Given these facts of the NUE formula, an indicator trait for breeding strategies (*i.e.,* MU) should not aim to identically portray NUE. Instead, a more precise definition of the goal trait seems necessary, which should subsequently be depicted by MU phenotypes. This novel goal trait might be defined as the following:

"The N efficient holobiont is that individual, which enables

- *the maximal ruminal utilisation of dietary N by microbial protein synthesis*
- *the maximal reutilisation of host-assimilated NH_3-N by urea recycling*
- *the balanced symbiotic interaction between host and rumen microbiota, which rebalances dynamically, when exposed to environmental changes (e.g.: diet)*
- *the maximal intestinal AA and (di-)peptide absorption*
- *the maximal utilisation of AA and (di-)peptides by the host for endogenous protein synthesis, reflected by body maintenance, growth, and milk protein yield*
- *the maintenance to remain healthy and fertile."*

8.4.2 MU as an indicator trait for the "N efficient holobiont"

MU data is available at a large scale. Moreover, MU is advantageous compared to the NUE formula due to its direct determination in milk, instead of a calculated ratio. However, considering MU as an indicator trait for a breeding strategy, which should depict the novel N efficiency trait, the following aspects of MU data need to be considered:

A major impact of the absolute N intake, and consequently of DMI intake on MU levels was described (*e.g.:* (Burgos *et al.*, 2007), chapter 3.3.4). Hence, cow-individual MU record data is assumed to be biased by cow-individual DMI variability. This assumption is substantiated by negative genetic correlations between MU and fertility parameters (*e.g.:* female fertility r= 0.15 in early lactation (Chen *et al.*, 2021a)), the latter being an indicator of energy balance and DMI as described above (*e.g.*: (Pryce *et al.*, 2014, Ferreira Júnior *et al.*, 2018), chapter 3.3.1). At this time point, this bias of DMI cannot be corrected due to the lack of (daily measured) cow-individual DMI data. Thus, although MU is available at a large scale, selecting for low MU may result in selecting for low DMI, thus may have crucial consequences for the energy balance of high yielding cows.

However, regarding that MU is a metabolite of detoxification instead of performance, the relationship of MU and DMI might be speculated as phenotypic correlation without genetic linkage. Although the

urea cycle is energy dependent (4 ATP per urea molecule), the absolute amount of energy costs for detoxification is, in comparison to body maintenance and milk performance (*i.e.,* protein synthesis), low (Von Engelhardt *et al.*, 2015). Basically, this would assume that breeding selection with MU is possible without negative effect on the energy balance, if the error of cow-individual DMI could precisely be quantified in a reference population. Regarding the high attractivity of efficiency traits for breeding selection, applicable solutions for cow-individual DMI measurements are already established on farm in Denmark, Sweden, and Finland (*e.g.:* camera techniques) and can be considered for high-scale application in the near future (Thomasen *et al.*, 2018, Seymour *et al.*, 2019). A correction of MU records via DMI data might therefore be approached. In this context, the following N specific aspects ought to be a matter for consideration:

The quantification of the actual N intake requires precise determination of absolute dietary CP including digestibility estimates of the respective N compounds (Fru Nji *et al.*, 2005). Furthermore, considering the separation of dietary compounds on the feeding table due to unhomogenised diets (Preißinger *et al.*, 2006), the calculated diets need to be validated by the actual provided diet on the feeding table by samples obtained on farm.

Enhanced DMI increases the dietary passage rate, which in turn lowers the absorption time for AA and (di-) peptides at the intestinal surface (Lebzien *et al.*, 2005). Considering that MU mainly origins from ruminal absorbed NPN (NH_3, NH^{4+} and urea) that passes the rumen epithelia via fast diffusion (Von Engelhardt *et al.*, 2015), the effect of an enhanced dietary passage rate on the total amount of host assimilated NPN and the resulting MU may only be minor. However, an increased dietary passage majorly decreases the dietary CP digestibility, due to the reduced time for rumen microbial proteolysis (Khol-Parisini and Zebeli, 2012). The reduced rumen microbial proteolysis in turn leads to less ruminal NPN (Riemeier, 2004), thus might consequently bias MU. If future techniques allow high scale cow-individual DMI data measurement, this bias might be quantified by MU measurements across different DMI levels.

The RNB is a proxy of ruminal N availability. In general, negative RNBs aim for the increased ruminal utilisation of recycled NPN by the microbes (Martens *et al.*, 2021). RNBs are calculated per kg DMI, thus accumulate in either ruminal-N over or N undersupply by each kg DMI variation, which in turn would lead to higher or lower ruminal NPN and thus potentially to higher or lower MU (Riemeier, 2004). In addition to the total amount of dietary CP and the digestibilities of the respective dietary N compounds, the RNB should therefore be additionally considered as a bias of DMI variation being present in MU record data.

8.4.3 MU for applied breeding selection: potential economic weight

The daily N intake relocates almost equally in milk, urine and faeces – diminished by minor amounts for maintenance requirements (*e.g.:* N intake: 498 g N/d; milk-N= 157 g N/d, urinary- N =159 g N/d, faecal-N= 180 g N/d in (Müller *et al.*, 2021)). In similarity to other quantification approaches of cow-individual N utilisation, MU focusses on the N retention in milk only. The phenotypic independency of faecal-N and MU was described above and further evidenced by the phenotypes in the Gut Dummerstorf herd. Breeding selection with MU is therefore assumed to not affect faecal-N. Considering faecal-N losses being majorly determined by the N digestion and N absorption capacity at the intestine (AA and (di-) peptides) a high similarity between faecal-N losses and feed efficiency can be assumed (Spiekers, 2021). Hence, a reduction of faecal-N losses may derive in tandem if feed efficiency or a similar trait is implemented in the breeding goal in the future.

In comparison to the small amount of urea in milk (*e.g.:* approx. 7% of total daily milk-N in (Müller *et al.*, 2021)), urea accounts for 70% to 80% of total urinary-N, resulting in absolute UU excretions of about 270g/d (Dijkstra *et al.*, 2013, Müller *et al.*, 2021). Thus, in dependency on the stability and validity of the relationship between MU and UU, a slight decrease in absolute MU would majorly impact absolute UU. Moreover, considering urea as a diuretic molecule, which influences the urine volume, a quantification of daily urine volumes in relation to MU and UU might be considered for the future.

Assuming that MU can be precisely differentiated from N intake in the future, and the MU to UU relationship exists, is stable and valid over selection (*i.e.,* genetically linked) an economic weight for a direct breeding value for MU would mainly refer to the quantification of UU yield. UU yield would display the majority of absolute urinary-N losses (see above: 70%-80% of urinary-N is urea-N) and accounts for approx. 35% of total N losses in urine and faeces (based on the examples above: urinary-N =159 g N/d, thereof on average 75% urea-N; faecal-N= 180 g N/d in (Müller *et al.*, 2021)). A combined calculation based on greenhouse gas savings as well as on a possible economic advantage due to feed N reduction could be suggested.

9 Conclusion

The integration of NUE into dairy cows' breeding programs was thought to be advantageous for socio-ecological, economic, and animal health's reasons. MU was suggested as indicator trait for NUE, due to its high cow-individual data availability, which enables genetic parameter and breeding value estimation in established pipelines. However, study approaches of the last years did not achieve a uniform knowledge on the stability and validity of the MU–NUE relationship. Moreover, cow-individual DMI was assumed to bias MU records. A deeper knowledge on underlying principles of cow-individual NUE is indispensable before MU can be seriously considered for breeding programs.

Due to the key function of the rumen microbiome in the N metabolism it was hypothesised that cow-individual NUE is determined by the entire "holobiont" (*i.e.,* host genome and rumen microbiome). Current knowledge on moderate to high heritabilities of specific rumen microbial taxa let assume that rumen microbial abundances are partly targeted by genomic enhanced breeding selection.

Four study approaches were conducted, which focused on the host genome–trait, the host genome–rumen microbiota and the entire host genome–rumen microbiota–trait axis. The studies aimed for a holistic exploration of cow-individual N utilisation and N excretion with focus on MU as a potential indicator trait for genomic enhanced breeding selection to increase NUE

Regarding the host genome–trait axis, genetic variability (narrowed down to a few candidate genes identified by SNP trait association) was identified, which is thought to influence the milk and urine N excretion phenotypes that were studied. Further analyses of these genomic regions may provide causal insights into the genetic-physiological relationships, leading to variability of renal and milk-N excretion in Holstein cows. Our analyses revealed no overlaps in significant regions between milk and urinary N excretion traits, which suggests an independent genetic architecture of the trait complexes. However, validation studies in larger cow populations are strongly suggested, particularly due to an overlapping QTL of MU and UU phenotypic raw models, without covariate-adjustment. If the genetic variant that uncovered the QTL is associated to a covariates variability, which biased the MU and UU raw models, or if it is indeed a genetic variant that associates with urea variability in both, milk and urine, asks for clarification. If the latter is the case, selecting for MU would potentially enable UU in- or decreases by pleiotropic effects.

The investigation of the host genome–rumen microbiota axis revealed that a specific pattern of microbial taxa and host transcript abundances distinguished the rumen profiles of Holsteins with high compared to low MU predisposition (HMug *vs.* LMUg). Host specific immune responses were evidenced to majorly determine this predisposed rumen microbial–host pattern. LMUg cows further adapted to the dietary N reduction by differentially regulated immune responses in the rumen tissue.

Although the underlying databases are based on human and mouse annotations it might be conceivable that breeding selection for MU target genomic regions which determine immune responses, thus influence maintenance and adaption of microbial taxa abundances with subsequent consequences for N utilisation and N excretion phenotypes.

Overlapping results of the rumen microbiota analyses in both, the Blue Cow and the Gut Dummerstorf cow population, implies that breeding selection for low MU might be accompanied by higher abundance of ureolytic and lower abundance of hyper-ammonia producing (HAB) rumen bacteria. This might enable an enhanced ruminal N utilisation and reduce N losses, due to a desirable ruminal (re) utilisation of urea and the absence of steep ruminal NH_3 increases and resulting effluxes in the host's NPN pool. Our analyses evidenced the *Chlostridia unclassified* genus (HAB bacterium) as the most validly determined ruminal player in the host genome–rumen microbiota–N associated trait axis. Our results would speculate lower ruminal abundance of *Chlostridia unclassified* as a consequence of an LMU host genotype, an interaction of this genus with the host specific ruminal immune response and a relation to LMU phenotypes by promoting the N incorporation into (milk) protein. However, this speculation is based on our statistical approaches, overlapping results between our studies and the current literature, thus is not causally proven. A validation and further exploration of our results in the context of the entire "holobiont" as a determinant of cow-individual N utilisation and N excretion may provide consistent results, which could subsequently be utilised for breeding application in the future.

Considering a serious application of GBVMU in breeding programs, it is mentionable that the phenotypes in the "Gut Dummerstorf" cow population revealed significantly lower MU, UU and slightly higher milk yields in GBV_{LMU} compared to GBV_{HMU} cows. Faecal-N concentrations, total milk-N and MP were not affected by GBVMU selection. Assuming that MU can be precisely differentiated from cow-individual N intake in the future, breeding selection with MU may enable greenhouse gas savings and a reduction of dietary N on dairy farms.

10 References

Althaus, B., Papke, G., and Sundrum, A. (2013). Use of near infrared reflectance spectroscopy to assess nitrogen and carbon fractions in dairy cow feces. *Animal Feed Science and Technology* 185(1-2):53-59.

Alves, K. L. G. C., Granja-Salcedo, Y. T., Messana, J. D., de Souza, V. C., Ganga, M. J. G., Colovate, P. H. D., Kishi, L. T., and Berchielli, T. T. (2020). Rumen bacterial diversity in relation to nitrogen retention in beef cattle. *Anaerobe*:102-316.

Aminlari, M. and Vaseghi, T. (1992). Arginase distribution in tissues of domestic animals. *Comparative Biochemistry and Physiology. B, Comparative Biochemistry* 103(2):385-389.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Nature Precedings*:1.

Ariyarathne, H. B. P. C., Correa-Luna, M., Blair, H. T., Garrick, D. J., and Lopez-Villalobos, N. (2021). Identification of genomic regions associated with concentrations of milk fat, protein, urea and efficiency of crude protein utilization in grazing dairy cows. *Genes* 12(3):456.

Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L., and Kraft, P. (2015). Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *The American Journal of Human Genetics* 96(2):329-339.

Atashi, H., Chen, Y., Wilmot, H., Vanderick, S., Hubin, X., and Gengler, N. (2022). Genome-wide association for milk urea concentration in Dual-Purpose Belgian Blue cows. *Journal of Animal Breeding and Genetics* 139(6):710-722.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., and Hu, T. T. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature* 465(7298):627-631.

Bach, A., Calsamiglia, S., and Stern, M. (2005). Nitrogen metabolism in the rumen. *Journal of Dairy science* 88:E9-E21.

Baker, G., Smith, J., and Cowan, D. (2003). Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* 55(3):541-555.

Beatson, P., Meier, S., Cullen, N., and Eding, H. (2019). Genetic variation in milk urea nitrogen concentration of dairy cattle and its implications for reducing urinary nitrogen excretion. *Animal* 13(10):2164-2171.

Beck, D., Jiang, H., and Zhang, L. (2009). Elucidating the evolutionary relationships among bos taurus digestive organs using unigene expression data. *International Journal of Evolutionary Biology*:Special section 2009:1-8.

Bell, C. G., Finer, S., Lindgren, C. M., Wilson, G. A., Rakyan, V. K., Teschendorff, A. E., Akan, P., Stupka, E., Down, T. A., and Prokopenko, I. (2010). Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PloS one* 5(11):e14040.

Bento, C. B. P., de Azevedo, A. C., Detmann, E., and Mantovani, H. C. (2015). Biochemical and genetic diversity of carbohydrate-fermenting and obligate amino acid-fermenting hyper-ammonia-producing bacteria from Nellore steers fed tropical forages and supplemented with casein. *BMC Microbiology* 15(1):1-15.

Bergamaschi, M., Maltecca, C., Schillebeeckx, C., McNulty, N. P., Schwab, C., Shull, C., Fix, J., and Tiezzi, F. (2020). Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Scientific Reports* 10(1):1-12.

Bergen, W., Purser, D., and Cline, J. (1968). Determination of limiting amino acids of rumen-isolated microbial proteins fed to rat. *Journal of Dairy Science* 51(10):1698-1700.

Berry, D. P. (2013). Breeding strategies to reduce environmental footprint in dairy cattle. *Advances in Animal Biosciences* 4(s1):28-36.

Bessman, N. J. and Sonnenberg, G. F. (2016). Emerging roles for antigen presentation in establishing host–microbiome symbiosis. *Immunological Reviews* 272(1):139-150.

Biorama (2022). Harnstoff. http://biorama.s3-website-eu-west-1.amazonaws.com/biblio/b90laka/lakah/har010.htm, access: 02.10.2022.

Bouwman, A., Schopen, G., Bovenhuis, H., Visker, M., and van Arendonk, J. (2010). Genome-wide scan to detect quantitative trait loci for milk urea nitrogen in Dutch Holstein-Friesian cows. *Journal of Dairy Science* 93(7):3310-3319.

Bristow, A. W., Whitehead, D. C., and Cockburn, J. E. (1992). Nitrogenous constituents in the urine of cattle, sheep and goats. *Journal of the Science of Food and Agriculture* 59(3):387-394.

Broderick, G. A. and Clayton, M. K. (1997). A statistical evaluation of animal and nutritional factors influencing concentrations of milk urea nitrogen. *Journal of Dairy Science* 80(11):2964-2971.

BRS, N. N. (2022). Rind-Deutsche Holsteins-Zuchtwertschätzung-Population, Geschichte, Genomische Selektion, Zuchtwertschätzung. Bundesverband Rind und Schwein e.V. (BRS), Adenauerallee 174, 53113 Bonn, https://www.rind-schwein.de/brs-rind/zuchtwertschaetzung-2.html, access: 29.09.2022.

Bryant, M. (1970). Normal flora—rumen bacteria. *The American Journal of Clinical Nutrition* 23(11):1440-1450.

Burgos, S. A., Fadel, J. G., and Depeters, E. J. (2007). Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: relation of milk urea nitrogen to urine urea nitrogen excretion. *Journal of Dairy Science* 90(12):5499-5508.

Buttchereit, N., Stamer, E., Junge, W., and Thaller, G. (2011). Genetic relationships among daily energy balance, feed intake, body condition score, and fat to protein ratio of milk in dairy cows. *Journal of Dairy Science* 94(3):1586-1591.

Calsamiglia, S., Ferret, A., Reynolds, C. K., Kristensen, N. B., and van Vuuren, A. M. (2010). Strategies for optimizing nitrogen use by ruminants. *Animal* 4(7):1184-1196.

Calus, M. P. and Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution* 43(1):1-14.

Cánovas, A., Rincon, G., Islas-Trejo, A., Wickramasinghe, S., and Medrano, J. F. (2010). SNP discovery in the bovine milk transcriptome using RNA-Seq technology. *Mammalian Genome* 21(11):592-598.

Cantalapiedra-Hijar, G., Dewhurst, R., Cheng, L., Cabrita, A., Fonseca, A., Fouillet, H., Noziere, P., and Marty, I. O. (2016). Isotopic N fractionation as a biomarker of nitrogen use efficiency by ruminants: A meta-analysis. *In Proc. of the 5. EAAP Symposium on Energy and Protein Metabolism and Nutrition.* Cracovie, Poland.

Cecchinato, A., Ribeca, C., Chessa, S., Cipolat-Gotet, C., Maretto, F., Casellas, J., and Bittante, G. (2014). Candidate gene association analysis for milk yield, composition, urea nitrogen and somatic cell scores in Brown Swiss cows. *Animal* 8(7):1062-1070.

Chen, Y., Atashi, H., Vanderick, S., Mota, R., Soyeurt, H., Hammami, H., and Gengler, N. (2021a). Genetic analysis of milk urea concentration and its genetic relationship with selected traits of interest in dairy cows. *Journal of Dairy Science* 104(12):12741-12755.

Chen, Y., Vanderick, S., Mota, R., Grelet, C., Gengler, N., and Consortium, G. (2021b). Estimation of genetic parameters for predicted nitrogen use efficiency and losses in early lactation of Holstein cows. *Journal of Dairy Science* 104(4):4413-4423.

Cheng, L., Woodward, S., Dewhurst, R. J., Zhou, H., and Edwards, G. (2014). Nitrogen partitioning, energy use efficiency and isotopic fractionation measurements from cows differing in genetic merit fed low-quality pasture in late lactation. *Animal Production Science* 54(10):1651-1656.

Cholewińska, P., Czyż, K., Nowakowski, P., and Wyrostek, A. (2020). The microbiome of the digestive system of ruminants–a review. *Animal Health Research Reviews* 21(1):3-14.

Cholewińska, P., Górniak, W., and Wojnarowski, K. (2021). Impact of selected environmental factors on microbiome of the digestive tract of ruminants. *BMC Veterinary Research* 17(1):1-10.

Colmenero, J. O. and Broderick, G. A. (2006). Effect of dietary crude protein concentration on milk production and nitrogen utilization in lactating dairy cows. *Journal of Dairy Science* 89(5):1704-1712.

Connor, E. E., Baldwin, R. L., Li, C.-j., Li, R. W., and Chung, H. (2013). Gene expression in bovine rumen epithelium during weaning identifies molecular regulators of rumen development and growth. *Functional & Integrative Genomics* 13(1):133-142.

Correa-Luna, M., Donaghy, D., Kemp, P., Schutz, M., and López-Villalobos, N. (2021). Nitrogen Use Efficiency and Excretion in Grazing Cows with High and Low Milk Urea Nitrogen Breeding Values. *Sustainability* 13(17):9827.

Cox, C. (2013). Experimentelle Untersuchungen zur Kinetik des Stickstoff-Umsatzes bei wachsenden Ziegenlämmern unter Verwendung des stabilen Isotops 15N. Dissertation. Fachbereich Veterinärmedizin, Justus Liebig University Gießen, Germany.

Curwen, V., Eyras, E., Andrews, T. D., Clarke, L., Mongin, E., Searle, S. M., and Clamp, M. (2004). The Ensembl automatic gene annotation system. *Genome Research* 14(5):942-950.

Dahlhoff, K. (2014). Beratung von milchviehhaltenden Betrieben auf der Basis von Verhaltens-und Erscheinungsparametern ihrer Milchkühe. Dissertation. Institut für Landtechnik, Landwirtschaftliche Fakultät, der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Difford, G. F., Løvendahl, P., Lassen, J., Guldbrandtsen, B., and Sahana, G. (2018a). Microbiability–new insights into (genetic) modelling methane emissions of cattle. *In Proc. of the World Congress on Genetics Applied to Livestock Production* (11): 405-410.Difford, G. F., Plichta, D. R., Løvendahl, P., Lassen, J., Noel, S. J., Højberg, O., Wright, A.-D. G., Zhu, Z., Kristensen, L., and Nielsen, H. B. (2018b). Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genetics* 14(10):e1007580.

Dijkstra, J., Oenema, O., Van Groenigen, J., Spek, J., Van Vuuren, A., and Bannink, A. (2013). Diet effects on urine composition of cattle and N2O emissions. *Animal* 7(s2):292-302.

Do, D. N., Ostersen, T., Strathe, A. B., Mark, T., Jensen, J., and Kadarmideen, H. N. (2014). Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs. *BMC Genetics* 15(1):1-15.

Ellis, S. A. and Ballingall, K. T. (1999). Cattle MHC: evolution in action? *Immunological Reviews* 167(1):159-168.

Elsik, C. G., Tellam, R. L., Worley, K. C., Gibbs, R. A., Muzny, D. M., Weinstock, G. M., Adelson, D. L., Eichler, E. E., and Elnitski, L. (2009). The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 324(5926):522-528.

Emmanuel, B. (1980). Urea cycle enzymes in tissues (liver, rumen epithelium, heart, kidney, lung and spleen) of sheep (Ovis aries). *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 65(4):693-697.

Ferreira Júnior, R. J., Bonilha, S. F., Monteiro, F. M., Cyrillo, J. N., Branco, R. H., Silva, J. A. I. V., and Mercadante, M. E. Z. (2018). Evidence of negative relationship between female fertility and feed efficiency in Nellore cattle. *Journal of Animal Science* 96(10):4035-4044.

Ferreira, M., Delagarde, R., and Edouard, N. (2021). CowNflow: A dataset on nitrogen flows and balances in dairy cows fed maize forage or herbage-based diets. *Data in Brief* 38:107393.

Flachowsky, G., Jeroch, H., Kirchgeßner, M., Pallauf, J., Pfeffer, E., Schulz, E., and Staudacher, W. (2001). Empfehlungen zur Energie- und Nährstoffversorgung der Milchkühe und Aufzuchtrinder. Book. Deutsche Landwirtschafts-Gesellschaft Verlag, Frankfurt am Main.

Fortes, M. R., Reverter, A., Zhang, Y., Collis, E., Nagaraj, S. H., Jonsson, N. N., Prayaga, K. C., Barris, W., and Hawken, R. J. (2010). Association weight matrix for the genetic dissection of puberty in beef cattle. *Proceedings of the National Academy of Sciences* 107(31):13642-13647.

Fru Nji, F., Hachenberg, S., and Pfeffer, E. (2005). Mikrobielle Proteinsynthese und Proteinabbau von Futtermitteln in einem modifizierten Hohenheimer Futterwerttest (HFT). Institut für Tierernährung, Landwirtschaftliche Fakultät, Rheinische Friedrich-Wilhelms-Universität Bonn, Lehr- und Forschungsschwerpunkt „Umweltverträgliche und Standortgerechte Landwirtschaft“, Forschungsbericht Nr. 125.

Giallongo, F., Hristov, A. N., Oh, J., Frederick, T., Weeks, H., Werner, J., Lapierre, H., Patton, R., Gehman, A., and Parys, C. (2015). Effects of slow-release urea and rumen-protected methionine and histidine on performance of dairy cows. *Journal of Dairy Science* 98(5):3292-3308.

Gibbs, R., Weinstock, G., Kappes, S., Schook, L., Skow, L., and Womack, J. (2002). Bovine genomic sequencing initiative. www.genome.gov/Pages/Research/Sequencing/SeqProposals/BovineSEQ.pdf.

Glatz-Hoppe, J. S. (2019). Weiterentwicklung von Bewertungswerkzeugen zur Charakterisierung der Versorgungssituation von Milchkühen mittels Daten aus der Milchkontrolle. Dissertation. Fachbereich Tiergesundheit und Tierschutz, Agrar-und Umweltwissenschaftliche Fakultät, Universität Rostock.

Gleeson, M., Nieman, D. C., and Pedersen, B. K. (2004). Exercise, nutrition and immune function. *Journal of Sports Sciences* 22(1):115-125.

Golder, H. M., Thomson, J. M., Denman, S. E., McSweeney, C. S., and Lean, I. J. (2018). Genetic markers are associated with the ruminal microbiome and metabolome in grain and sugar challenged dairy heifers. *Frontiers in Genetics* 9:62.

Gonda, H. L. and Lindberg, J. E. (1994). Evaluation of dietary nitrogen utilization in dairy cows based on urea concentrations in blood, urine and milk, and on urinary concentration of purine derivatives. *Acta Agriculturae Scandinavica A-Animal Sciences* 44(4):236-245.

Gondret, F., Vincent, A., Houée-Bigot, M., Siegel, A., Lagarrigue, S., Causeur, D., Gilbert, H., and Louveau, I. (2017). A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. *BMC Genomics* 18(1):1-17.

Graham, C. and Simmons, N. L. (2005). Functional organization of the bovine rumen epithelium. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology* 288(1):R173-181.

Grelet, C., Froidmont, E., Foldager, L., Salavati, M., Hostens, M., Ferris, C. P., Ingvartsen, K. L., Crowe, M. A., Sorensen, M. T., and Pierna, J. F. (2020). Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. *Journal of Dairy Science* 103(5):4435-4445.

Grieve, D., Korver, S., Rijpkema, Y., and Hof, G. (1986). Relationship between milk composition and some nutritional parameters in early lactation. *Livestock Production Science* 14(3):239-254.

Groenestein, C., Hutchings, N., Haenel, H., Amon, B., Menzi, H., Mikkelsen, M., Misselbrook, T., Van Bruggen, C., Kupper, T., and Webb, J. (2019). Comparison of ammonia emissions related to nitrogen use efficiency of livestock production in Europe. *Journal of Cleaner Production* 211:1162-1170.

Gruber, L., Pries, M., Schwarz, F., Spiekers, H., and Staudacher, W. (2006). Schätzung der Futteraufnahme bei der Milchkuh. *DLG-Information* 1(2006):1-29.

Guarini, A. R., Lourenco, D. A. L., Brito, L. F., Sargolzaei, M., Baes, C. F., Miglior, F., Misztal, I., and Schenkel, F. S. (2019). Genetics and genomics of reproductive disorders in Canadian Holstein cattle. *Journal of Dairy Science* 102(2):1341-1353.

Guliński, P., Salamończyk, E., and Młynek, K. (2016). Improving nitrogen use efficiency of dairy cows in relation to urea in milk–a review. *Animal Science Papers and Reports* 34(1):5-24.

Halushka, M. K., Fan, J.-B., Bentley, K., Hsie, L., Shen, N., Weder, A., Cooper, R., Lipshutz, R., and Chakravarti, A. (1999). Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nature Genetics* 22(3):239-247.

Hartinger, T., Gresner, N., and Südekum, K.-H. (2018). Does intra-ruminal nitrogen recycling waste valuable resources? A review of major players and their manipulation. *Journal of Animal Science and Biotechnology* 9(1):1-21.

Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science* 92(2):433-443.

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., and Janssen, P. H. (2015). Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Scientific Reports* 5(1):1-15.

Heuer, C., Van Straalen, W., Schukken, Y., Dirkzwager, A., and Noordhuizen, J. (2000). Prediction of energy balance in a high yielding dairy herd in early lactation: model development and precision. *Livestock Production Science* 65(1-2):91-105.

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., and Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences* 106(23):9362-9367.

Hof, G., Vervoorn, M., Lenaers, P., and Tamminga, S. (1997). Milk urea nitrogen as a tool to monitor the protein nutrition of dairy cows. *Journal of Dairy Science* 80(12):3333-3340.

Holmans, P. (2010). Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. *Advances in Genetics* 72:141-179.

Höltje, C. (2018). Neues bei der Zuchtwertschätzung: Zuchtwerte für Klauen und Gesundheitsmerkmale. Presentation in *Haus Düsse, Fruchtbarkeitsseminar, December 2018,*https://www.ruweg.de/fileadmin/user_upload/aktuelles/News/News_2018/Fruchtbarkeitsseminar/Fruchtbarkeitsseminar_Haus_Duesse_Hoeltje_20181206.pdf.

Honerlagen, H., Reyer, H., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., Reinsch, N., and Wimmers, K. (2021). Identification of Genomic Regions Influencing N-Metabolism and N-Excretion in Lactating Holstein-Friesians. *Frontiers in Genetics* 12:1241.

Honerlagen, H., Reyer, H., Segelke, D., Müller, C. B. M., Prahl, M. C., Ponsuksili, S., Trakooljul, N., Reinsch, N., Kuhla, B., and Wimmers, K. (2022). Ruminal background of predisposed milk urea (MU) concentration in Holsteins. *Frontiers in Microbiology* 13.

Houlahan, K., Schenkel, F. S., Hailemariam, D., Lassen, J., Kargo, M., Cole, J. B., Connor, E. E., Wegmann, S., Oliveira Jr, G. A., and Miglior, F. (2021). Effects of incorporating dry matter intake and residual feed intake into a selection index for dairy cattle using deterministic modeling. *Animals* 11(4):1157.

Hu, Y., Chen, I.-P., de Almeida, S., Tiziani, V., Do Amaral, C. M. R., Gowrishankar, K., Passos-Bueno, M. R., and Reichenberger, E. J. (2013). A novel autosomal recessive GJA1 missense mutation linked to Craniometaphyseal dysplasia. *PloS one* 8(8):e73576.

Huang, Q. (2015). Genetic study of complex diseases in the post-GWAS era. *Journal of Genetics and Genomics* 42(3):87-98.

Huhtanen, P. and Hristov, A. N. (2009). A meta-analysis of the effects of dietary protein concentration and degradability on milk protein yield and milk N efficiency in dairy cows. *Journal of Dairy Science* 92(7):3222-3232.

Huth, A. (2012). Studien zum mRNA profiling an humanem postmortalem Gewebe. Dissertation. Naturwissenschaftliche Fakultät, Julius-Maximilians-Universität Würzburg.

Ibal, J. C., Park, Y.-J., Park, M.-K., Lee, J., Kim, M.-C., and Shin, J.-H. (2022). Review of the Current State of Freely Accessible Web Tools for the Analysis of 16S rRNA Sequencing of the Gut Microbiome. *International Journal of Molecular Sciences* 23(18):10865.

Jahnel, R. E., Wittenburg, D., Mayer, M., Taeubert, H., and Reinsch, N. (2021). Estimation of genetic parameters of milk urea content. *Züchtungskunde* 93(2):157-168.

Jehl, F., Degalez, F., Bernard, M., Lecerf, F., Lagoutte, L., Désert, C., Coulée, M., Bouchez, O., Leroux, S., and Abasht, B. (2021). RNA-seq data for reliable SNP detection and genotype calling: interest for coding variant characterization and cis-regulation analysis by allele-specific expression in livestock species. *Frontiers in Genetics* 12:1104.

Jentsch, T. J., Stein, V., Weinreich, F., and Zdebik, A. A. (2002). Molecular structure and physiological function of chloride channels. *Physiological Reviews* 82(2):503-568.

Jin, D., Zhao, S., Zheng, N., Beckers, Y., and Wang, J. (2018). Urea metabolism and regulation by rumen bacterial urease in ruminants–a review. *Annals of Animal Science* 18(2):303.

Jonker, J., Kohn, R., and High, J. (2002). Dairy herd management practices that impact nitrogen utilization efficiency. *Journal of Dairy Science* 85(5):1218-1226.

Kand, D. and Dickhoefer, U. (2021). The effects of rumen nitrogen balance on nutrient intake, nitrogen partitioning, and microbial protein synthesis in lactating dairy cows offered different dietary protein sources. *Journal of Dairy Science* 104(4):4223-4235.

Karisa, B., Moore, S., and Plastow, G. (2013). Complicated relationships: a review of biological interaction networks and pathways in animal science. *Springer Science Reviews* 1(1):73-83.

Kauffman, A. and St-Pierre, N. (2001). The relationship of milk urea nitrogen to urine nitrogen excretion in Holstein and Jersey cows. *Journal of Dairy Science* 84(10):2284-2294.

Kebreab, E., Strathe, A., Dijkstra, J., Mills, J. A., Reynolds, C. K., Crompton, L. A., Yan, T., and France, J. (2010). Energy and protein interactions and their effect on nitrogen excretion in dairy cows. *In Symp. on Energy and Protein Metabolism and Nutrition*, Parma, Italy: 417-426.

Khafipour, E., Li, S., Tun, H., Derakhshani, H., Moossavi, S., and Plaizier, J. (2016). Effects of grain feeding on microbiota in the digestive tract of cattle. *Animal Frontiers* 6(2):13-19.

Khol-Parisini, A. and Zebeli, Q. (2012). Die Milchkuh im Spannungsfeld zwischen Leistung, Gesundheit und Nährstoffeffizienz.*In: 39. Viehwirtschaftliche Fachtagung*. Lehr-und Forschungszentrum für Landwirtschaft Raumberg-Gumpenstein, Germany:43-50

Kim, N. H., Kim, Y. J., Oh, J. H., and Cho, Y. S. (2012). Replication of the Association of the 6q22. 31c Locus near GJA1 with Pulse Rate in the Korean Population. *Genomics & Informatics* 10(2):106.

Kohn, R., Dinneen, M., and Russek-Cohen, E. (2005). Using blood urea nitrogen to predict nitrogen excretion and efficiency of nitrogen utilization in cattle, sheep, goats, horses, pigs, and rats. *Journal of Animal Science* 83(4):879-889.

König, S., Chang, Y., Borstel, U., Gianola, D., and Simianer, H. (2008). Genetic and phenotypic relationships among milk urea nitrogen, fertility, and milk yield in Holstein cows. *Journal of Dairy Science* 91(11):4372-4382.

Kool, D. M., Hoffland, E., Hummelink, E. W., and Van Groenigen, J. W. (2006). Increased hippuric acid content of urine can reduce soil N2O fluxes. *Soil Biology and Biochemistry* 38(5):1021-1027.

Korte, A. and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9(1):1-9.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology* 79(17):5112-5120.

Kraft, W. (2005). Klinische Labordiagnostik in der Tiermedizin. Vol. 7. Schattauer Verlag, J. G. Cotta'sche Buchhandlung Nachfolger GmbH, Stuttgart/ New York.

Krpálková, L., O'Mahony, N., Carvalho, A., Campbell, S., Corkery, G., Broderick, E., Riordan, D., and Walsh, J. (2021). Association of Feed Efficiency, Feeding Rate, and Behaviour with the Milk Performance of Dairy Cows. *Dairy* 2(4):684-694.

Kvidera, S., Horst, E., Abuajamieh, M., Mayorga, E., Fernandez, M. S., and Baumgard, L. (2017). Glucose requirements of an activated immune system in lactating Holstein cows. *Journal of Dairy Science* 100(3):2360-2374.

Lahart, B., McParland, S., Kennedy, E., Boland, T., Condon, T., Williams, M., Galvin, N., McCarthy, B., and Buckley, F. (2019). Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis. *Journal of Dairy Science* 102(10):8907-8918.

Lamont, S., Hou, Y.-H., Young, B., and Nordskog, A. (1987). Research note: differences in major histocompatibility complex gene frequencies associated with feed efficiency and laying performance. *Poultry Science* 66(6):1064-1066.

Lavery, A. and Ferris, C. P. (2021). Proxy Measures and Novel Strategies for Estimating Nitrogen Utilisation Efficiency in Dairy Cattle. *Animals* 11(2):343.

Lebzien, P., Meyer, U., Flachowsky, G., and von Erbsen, V. D. E. (2005). Ernährung und Fütterung des Rindes : 6.1, Ernährungsphysiologische Grundlagen. *Landbauforschung Völkenrode Sonderheft* 289: 89-101.

Lee, C., Hristov, A., Heyler, K., Cassidy, T., Long, M., Corl, B., and Karnati, S. (2011). Effects of dietary protein concentration and coconut oil supplementation on nitrogen utilization and production in dairy cows. *Journal of Dairy Science* 94(11):5544-5557.

Lewin, H. A., Russell, G. C., and Glass, E. J. (1999). Comparative organization and function of the major histocompatibility complex of domesticated cattle. *Immunological Reviews* 167(1):145-158.

LfL. (2013). Eiweißfuttermittel in der Rinderfütterung. *In LfL-Information, 2. Auflage.* Bayerische Landesanstalt für Landwirtschaft (LfL), Vöttinger Straße 38, 85354 Freising-Weihenstephan.

Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., Fitzsimmons, C., and Plastow, G. (2019). Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. *Microbiome* 7(1):1-17.

Libera, K., Szumacher-Strabel, M., Vazirigohar, M., Zieliński, W., Lukow, R., Wysocka, K., Kołodziejski, P., Lechniak, D., Varadyova, Z., and Patra, A. K. (2021). Effects of feeding urea-treated triticale and oat grain mixtures on ruminal fermentation, microbial population, and milk production performance of midlactation dairy cows. *Annals of Animal Science* 21(3):1007-1025.

Lindberg, J. (1989). Nitrogen metabolism and urinary excretion of purines in goat kids. *British Journal of Nutrition* 61(2):309-321.

Liu, K., Zhang, Y., Yu, Z., Xu, Q., Zheng, N., Zhao, S., Huang, G., and Wang, J. (2021). Ruminal microbiota–host interaction and its effect on nutrient metabolism. *Animal Nutrition* 7(1):49-55.

Liu, Z., Seefried, F., Reinhardt, F., Thaller, G., and Reents, R. (2010). Dairy cattle genetic evaluation enhanced with genomic information. *In Proc. of the 9th WCGALP,* Leipzig, Germany.

Lloréns-Rico, V., Vieira-Silva, S., Gonçalves, P. J., Falony, G., and Raes, J. (2021). Benchmarking microbiome transformations favors experimental quantitative approaches to address compositionality and sampling depth biases. *Nature Communications* 12(1):1-12.

Loor, J. J., Elolimy, A. A., and McCann, J. C. (2016). Dietary impacts on rumen microbiota in beef and dairy production. *Animal Frontiers* 6(3):22-29.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12):1-21.

Maas, J., France, J., and McBride, B. (1997). Model of milk protein synthesis. A mechanistic model of milk protein synthesis in the lactating bovine mammary gland. *Journal of Theoretical Biology* 187(3):363-378.

Mariotti, F., Tomé, D., and Mirand, P. P. (2008). Converting nitrogen into protein—beyond 6.25 and Jones' factors. *Critical Reviews in Food Science and Nutrition* 48(2):177-184.

Martens, S., Steinhöfel, O., and Richardt, W. (2021). Umweltgerechte Ernährung von Milchrindern. Schriftenreihe 17, Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG), Pillnitzer Platz 3, 01326 Dresden.

Martínez-Álvaro, M., Auffret, M. D., Duthie, C.-A., Dewhurst, R. J., Cleveland, M. A., Watson, M., and Roehe, R. (2022). Bovine host genome acts on rumen microbiome function linked to methane emissions. *Communications Biology* 5(1):1-16.

Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sandstrom, R., Qu, H., and Brody, J. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337(6099):1190-1195.

McCabe, M. S., Cormican, P., Keogh, K., O'Connor, A., O'Hara, E., Palladino, R. A., Kenny, D. A., and Waters, S. M. (2015). Illumina MiSeq phylogenetic amplicon sequencing shows a large reduction of an uncharacterised Succinivibrionaceae and an increase of the Methanobrevibacter gottschalkii clade in feed restricted cattle. *PloS one* 10(7):e0133234.

McCann, J. C., Elolimy, A. A., and Loor, J. J. (2017). Rumen microbiome, probiotics, and fermentation additives. *Veterinary Clinics: Food Animal Practice* 33(3):539-553.

McCann, J. C., Wickersham, T. A., and Loor, J. J. (2014). High-throughput methods redefine the rumen microbiome and its relationship with nutrition and metabolism. *Bioinformatics and Biology Insights* 8:153-189.

McLoughlin, K. E., Nalpas, N. C., Rue-Albrecht, K., Browne, J. A., Magee, D. A., Killick, K. E., Park, S. D., Hokamp, K., Meade, K. G., and O'Farrelly, C. (2014). RNA-seq transcriptional profiling of peripheral blood leukocytes from cattle infected with Mycobacterium bovis. *Frontiers in Immunology* 5:396.

Meuwissen, T. H., Hayes, B. J., and Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4):1819-1829.

Miglior, F., Sewalem, A., Jamrozik, J., Bohmanova, J., Lefebvre, D. M., and Moore, R. K. (2007). Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in Canadian Holstein cattle. *Journal of Dairy Science* 90(5):2468-2479.

Müller, C. B. M., Görs, S., Derno, M., Tuchscherer, A., Wimmers, K., Zeyner, A., and Kuhla, B. (2021). Differences between Holstein dairy cows in renal clearance rate of urea affect milk urea concentration and the relationship between milk urea and urinary nitrogen excretion. *Science of the Total Environment* 755(2):143-198.

Munyaneza, N., Niyukuri, J., and Hachimi, Y. E. (2017). Milk Urea Nitrogen as an Indicator of Nitrogen Metabolism Efficiency in Dairy Cows: A Review. *Theriogenology Insight* 7(3):145-159.

Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N. E., Ahfeldt, T., Sachs, K. V., Li, X., Li, H., Kuperwasser, N., and Ruda, V. M. (2010). From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 466(7307):714-719.

Mutsvangwa, T., Davies, K., McKinnon, J., and Christensen, D. (2016). Effects of dietary crude protein and rumen-degradable protein concentrations on urea recycling, nitrogen balance, omasal nutrient flow, and milk production in dairy cows. *Journal of Dairy Science* 99(8):6298-6310.

Myer, P. R., Smith, T. P., Wells, J. E., Kuehn, L. A., and Freetly, H. C. (2015). Rumen microbiome from steers differing in feed efficiency. *PloS one* 10(6):e0129174.

Neyrinck, A. M., Possemiers, S., Druart, C., Van de Wiele, T., De Backer, F., Cani, P. D., Larondelle, Y., and Delzenne, N. M. (2011). Prebiotic effects of wheat arabinoxylan related to the increase in bifidobacteria, Roseburia and Bacteroides/Prevotella in diet-induced obese mice. *PloS one* 6(6):e20944.

Notebaart, R. A., Van Enckevort, F. H., Francke, C., Siezen, R. J., and Teusink, B. (2006). Accelerating the reconstruction of genome-scale metabolic networks. *BMC Bioinformatics* 7(1):1-10.

Nousiainen, J., Shingfield, K., and Huhtanen, P. (2004). Evaluation of milk urea nitrogen as a diagnostic of protein feeding. *Journal of Dairy Science* 87(2):386-398.

Paster, B. J., Russell, J. B., Yang, C., Chow, J., Woese, C. R., and Tanner, R. (1993). Phylogeny of the ammonia-producing ruminal bacteria Peptostreptococcus anaerobius, Clostridium sticklandii, and Clostridium aminophilum sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 43(1):107-110.

Pegolo, S., Dadousis, C., Mach, N., Ramayo-Caldas, Y., Mele, M., Conte, G., Schiavon, S., Bittante, G., and Cecchinato, A. (2017). SNP co-association and network analyses identify E2F3, KDM5A and BACH2 as key regulators of the bovine milk fatty acid profile. *Scientific Reports* 7(1):17317.

Pérez-Enciso, M., Zingaretti, L. M., Ramayo-Caldas, Y., and de Los Campos, G. (2021). Opportunities and limits of combining microbiome and genome data for complex trait prediction. *Genetics Selection Evolution* 53(1):1-20.

Peterson, D. A., McNulty, N. P., Guruge, J. L., and Gordon, J. I. (2007). IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell Host & Microbe* 2(5):328-339.

Powell, J., Gourley, C., Rotz, C., and Weaver, D. M. (2010). Nitrogen use efficiency: A potential performance indicator and policy tool for dairy farms. *Environmental Science & Policy* 13(3):217-228.

Powell, J., Jackson-Smith, D. B., McCrory, D., Saam, H., and Mariola, M. (2006). Validation of feed and manure data collected on Wisconsin dairy farms. *Journal of Dairy Science* 89(6):2268-2278.

Preißinger, W., Obermaier, A., and Spiekers, H. (2006). Auswirkung gestaffelter Strukturwerte (Häcksellänge von Maissilage, Kraftfutterzusammensetzung) bei der Milchkuh. *In Forum angewandte Forschung in der Rinder-und Schweinefütterung*, Fulda, Germany :35-41.

Pryce, J., Wales, W. d., De Haas, Y., Veerkamp, R., and Hayes, B. (2014). Genomic selection for feed efficiency in dairy cattle. *Animal* 8(1):1-10.

Rafiq, S., Huma, N., Pasha, I., Sameen, A., Mukhtar, O., and Khan, M. I. (2016). Chemical Composition, Nitrogen Fractions and Amino Acids Profile of Milk from Different Animal Species. *Asian-Australasian Journal of Animal Sciences* 29(7):1022-1028.

Ramayo-Caldas, Y., Zingaretti, L., Popova, M., Estellé, J., Bernard, A., Pons, N., Bellot, P., Mach, N., Rau, A., and Roume, H. (2020). Identification of rumen microbial biomarkers linked to methane emission in Holstein dairy cows. *Journal of Animal Breeding and Genetics* 137(1):49-59.

Razmi, N., Ali Jelodar, G., Nazifi, S., and Dehghani, A. (2005). Arginase status in cattle reproductive system. *Veterinarski arhiv* 75(1):31-38.

Reinhardt, F., Liu, Z., Seefried, F., Rensing, S., and Reents, R. (2011). Genomische Selektion: Stand der Implementierung bei Deutschen Holsteins. *Züchtungskunde* 83(4/5):248-256.

Ribeiro, G., Cesar, A. S. M., and Fukumasu, H. (2021). The search for genetic functional variants of feed efficiency in Nelore cattle. Dissertation. Faculty of Zootechny and Food Engineering, University of São Paulo.

Ribeiro, G. O., Oss, D. B., He, Z., Gruninger, R. J., Elekwachi, C., Forster, R. J., Yang, W., Beauchemin, K. A., and McAllister, T. A. (2017). Repeated inoculation of cattle rumen with bison rumen contents alters the rumen microbiome and improves nitrogen digestibility in cattle. *Scientific Reports* 7(1):1-16.

Riemeier, A. (2004). Einfluss der ruminalen Stickstoffbilanz (RNB) auf die Pansenfermentation, mikrobielle Proteinsynthese, Menge des am Dünndarm anflutenden nutzbaren Proteins (nXP) sowie der Stickstoffausscheidung. Dissertation. Institut für Tierernährung der Tierärztlichen Hochschule Hannover und Institut für Tierernährung der Bundesforschungsanstalt für Landwirtschaft in Braunschweig.

Roehe, R., Dewhurst, R. J., Duthie, C.-A., Rooke, J. A., McKain, N., Ross, D. W., Hyslop, J. J., Waterhouse, A., Freeman, T. C., and Watson, M. (2016). Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane emitting and efficiently feed converting hosts based on metagenomic gene abundance. *PLoS Genetics* 12(2):e1005846.

Saborío-Montero, A., Gutiérrez-Rivas, M., López-García, A., García-Rodríguez, A., Atxaerandio, R., Goiri, I., Jiménez-Montero, J. A., and González-Recio, O. (2021). Holobiont effect accounts for more methane emission variance than the additive and microbiome effects on dairy cattle. *Livestock Science* 250:104538.

Sasson, G., Kruger Ben-Shabat, S., Seroussi, E., Doron-Faigenboim, A., Shterzer, N., Yaacoby, S., Berg Miller, M. E., White, B. A., Halperin, E., and Mizrahi, I. (2017). Heritable bovine rumen bacteria are phylogenetically related and correlated with the cow's capacity to harvest energy from its feed. *MBio* 8(4):e00703-00717.

Schiavon, S., Cesaro, G., Cecchinato, A., Cipolat-Gotet, C., Tagliapietra, F., and Bittante, G. (2016). The influence of dietary nitrogen reduction and conjugated linoleic acid supply to dairy cows on fatty acids in milk and their transfer to ripened cheese. *Journal of Dairy Science* 99(11):8759-8778.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., and Robinson, C. J. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75(23):7537-7541.

Scholey, R., Evans, N., Blowey, R., Massey, J., Murray, R., Smith, R., Ollier, W., and Carter, S. (2013). Identifying host pathogenic pathways in bovine digital dermatitis by RNA-Seq analysis. *The Veterinary Journal* 197(3):699-706.

Schubert, N., Nichols, H. J., and Winternitz, J. C. (2021). How can the MHC mediate social odor via the microbiota community? A deep dive into mechanisms. *Behavioral Ecology* 32(3):359-373.

Schwertl, M., Auerswald, K., Schäufele, R., and Schnyder, H. (2005). Carbon and nitrogen stable isotope composition of cattle hair: ecological fingerprints of production systems? *Agriculture, Ecosystems & Environment* 109(1-2):153-165.

Schwiebert, E. M., Lopes, A. G., and Guggino, W. B. (1994). Chloride channels along the nephron. *Current Topis in Membranes* 42:265-315.

Sebek, L., Van Riel, J., and De Jong, G. (2007). The breeding value for milk urea as predictor for the efficiency of protein utilization in dairy cows. Rapport 81. Animal Sciences Group van Wageningen UR; Postbus 65, 8200 AB Lelystad, the Netherlands..

Seefried, F., Liu, Z., Thaller, G., and Reinhardt, F. (2010). Die genomische Zuchtwertschätzung bei der Rasse Deutsche Holstein. *Züchtungskunde* 1:14-21.

Seymour, D., Cánovas, A., Baes, C. F., Chud, T., Osborne, V., Cant, J., Brito, L., Gredler-Grandl, B., Finocchiaro, R., and Veerkamp, R. (2019). Invited review: Determination of large-scale individual dry matter intake phenotypes in dairy cattle. *Journal of Dairy Science* 102(9):7655-7663.

Shabat, S. K. B., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Berg Miller, M. E., White, B. A., Shterzer, N., and Mizrahi, I. (2016). Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *The ISME Journal* 10(12):2958-2972.

Sharif, S., Mallard, B., Wilkie, B., Sargeant, J., Scott, H., Dekkers, J., and Leslie, K. (1999). Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) with production traits in Canadian dairy cattle. *Animal Genetics* 30(2):157-160.

Shaw, J. and Petersen, W. (1938). Arginase in the mammary gland. *Proceedings of the Society for Experimental Biology and Medicine* 38(5):631-632.

Shetty, N., Løvendahl, P., Lund, M., and Buitenhuis, A. (2017). Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. *Journal of Dairy Science* 100(1):253-264.

Shook, G. (2006). Major advances in determining appropriate selection goals. *Journal of Dairy Science* 89(4):1349-1361.

Sinclair, K., Garnsworthy, P., Mann, G., and Sinclair, L. (2014). Reducing dietary protein in dairy cow diets: Implications for nitrogen utilization, milk production, welfare and fertility. *Animal* 8(2):262-274.

Small, K. S., Hedman, Å. K., Grundberg, E., Nica, A. C., Thorleifsson, G., Kong, A., Thorsteindottir, U., Shin, S.-Y., Richards, H. B., and Soranzo, N. (2011). Correction: Corrigendum: Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature Genetics* 43(10):1040-1040.

Spanghero, M. and Kowalski, Z. (1997). Critical analysis of N balance experiments with lactating cows. *Livestock Production Science* 52(2):113-122.

Spiekers, H., Hertel-Böhnke, P., Pfanzelt, S., Schiefler, I. (2021). Die optimale Kuh: gesund, effizient, umweltgerecht. Bayerische Landesanstalt für Landwirtschaft (LfL), Vöttinger Straße 38, 85354 Freising-Weihenstephan. Schriftenreihe ISSN 1611-4159.

Stains, J. P., Watkins, M. P., Grimston, S. K., Hebert, C., and Civitelli, R. (2014). Molecular mechanisms of osteoblast/osteocyte regulation by connexin43. *Calcified Tissue International* 94(1):55-67.

Stallcup, O., Davis, G., and Shields, L. (1975). Influence of dry matter and nitrogen intakes on fecal nitrogen losses in cattle. *Journal of Dairy Science* 58(9):1301-1307.

Stamer, E., Brade, W., Junge, W., and Thaller, G. (2009). Modellentwicklung und Schätzung genetischer Parameter für den Harnstoffgehalt in der Milch bei erstlaktierenden Holstein-Kühen. *Züchtungskunde* 81:361-374.

Stamer, E., Brade, W., and Thaller, G. (2011). Modellentwicklung und Schätzung genetischer Parameter für den Harnstoffgehalt in der Milch erst-und zweitlaktierender Holstein-Kühe. *Züchtungskunde. Bd* 83:104-117.

Stevenson, K. R., Coolon, J. D., and Wittkopp, P. J. (2013). Sources of bias in measures of allele-specific expression derived from RNA-seq data aligned to a single reference genome. *BMC Genomics* 14(1):1-13.

Stoop, W. M., Bovenhuis, H., and van Arendonk, J. A. (2007). Genetic parameters for milk urea nitrogen in relation to milk production traits. *Journal of Dairy Science* 90(4):1981-1986.

Tamanai-Shacoori, Z., Smida, I., Bousarghin, L., Loreal, O., Meuric, V., Fong, S. B., Bonnaure-Mallet, M., and Jolivet-Gougeon, A. (2017). Roseburia spp.: a marker of health? *Future Microbiology* 12(2):157-170.

Tan, B., Luo, W., Shen, Z., Xiao, M., Wu, S., Meng, X., Wu, X., Yang, Z., Tian, L., and Wang, X. (2019). Roseburia intestinalis inhibits oncostatin M and maintains tight junction integrity in a murine model of acute experimental colitis. *Scandinavian Journal of Gastroenterology* 54(4):432-440.

Tan, P., Liu, H., Zhao, J., Gu, X., Wei, X., Zhang, X., Ma, N., Johnston, L. J., Bai, Y., and Zhang, W. (2021). Amino acids metabolism by rumen microorganisms: Nutrition and ecology strategies to reduce nitrogen emissions from the inside to the outside. *Science of the Total Environment* 800:149-596.

Tannock, G. W. and Liu, Y. (2020). Guided dietary fibre intake as a means of directing short-chain fatty acid production by the gut microbiota. *Journal of the Royal Society of New Zealand* 50(3):434-455.

Thomasen, J. R., Lassen, J., Nielsen, G. G. B., Borggard, C., Stentebjerg, P. R. B., Hansen, R. H., Hansen, N. W., and Borchersen, S. (2018). Individual cow identification in a commercial herd using 3D camera technology. *In: In Proc. of the World Congress on Genetics Applied to Livestock Production, v. Technologies-Novel Phenotypes,* Auckland, New Zealand:11-16.

Tiefenthaller, F. (2009). Futterwerttabelle für Wiederkäuer-Grundfutter-Kraftfutter, Zusammenstellung-Stand 2016-09. Referat Fütterung, Abteilung Tierproduktion, Landwirtschaftkammer Oberösterreich, Auf der Gugl 3, 4021 Linz.

Timonen, A. A., Katholm, J., Petersen, A., Mõtus, K., and Kalmus, P. (2017). Within-herd prevalence of intramammary infection caused by Mycoplasma bovis and associations between cow udder health, milk yield, and composition. *Journal of Dairy Science* 100(8):6554-6561.

Turpin, W., Espin-Garcia, O., Xu, W., Silverberg, M. S., Kevans, D., Smith, M. I., Guttman, D. S., Griffiths, A., Panaccione, R., and Otley, A. (2016). Association of host genome with intestinal microbial composition in a large healthy cohort. *Nature Genetics* 48(11):1413-1417.

Uffelmann, E., Huang, Q. Q., Munung, N. S., De Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., and Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers* 1(1):1-21.

Uwizeye, A., de Boer, I. J. M., Opio, C. I., Schulte, R. P. O., Falcucci, A., Tempio, G., Teillard, F., Casu, F., Rulli, M., Galloway, J. N., Leip, A., Erisman, J. W., Robinson, T. P., Steinfeld, H., and Gerber, P. J. (2020). Nitrogen emissions along global livestock supply chains. *Nature Food* 1(7):437-446.

Valente, E., Araujo, L., Carvalho, S., and Stahlhofer, M. (2018). Breath ammonia as a bioindicator of protein nutrition in heifers. *Livestock Science* 218:97-100.

Van den Berg, I., Ho, P. N., Nguyen, T. V., Haile-Mariam, M., MacLeod, I. M., Beatson, P. R., O'Connor, E., and Pryce, J. E. (2022). GWAS and genomic prediction of milk urea nitrogen in Australian and New Zealand dairy cattle. *Genetics Selection Evolution* 54(1):1-12.

VandeHaar, M. J., Armentano, L. E., Weigel, K., Spurlock, D. M., Tempelman, R. J., and Veerkamp, R. (2016). Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. *Journal of Dairy Science* 99(6):4941-4954.

Vansteelandt, S., Goetgeluk, S., Lutz, S., Waldman, I., Lyon, H., Schadt, E. E., Weiss, S. T., and Lange, C. (2009). On the adjustment for covariates in genetic association analysis: a novel, simple principle to infer direct causal effects. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society* 33(5):394-405.

Vekaria, R. M., Shirley, D. G., Sevigny, J., and Unwin, R. J. (2006). Immunolocalization of ectonucleotidases along the rat nephron. *American Journal of Physiology-Renal Physiology* 290(2):F550-560.

Virtanen, A. I. (1966). Milk Production of Cows on Protein-Free Feed: Studies of the use of urea and ammonium salts as the sole nitrogen source open new important perspectives. *Science* 153(3744):1603-1614.

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., and Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics* 101(1):5-22.

Vit. (2022). Beschreibung der Zuchtwertschätzung für alle Schätzmerkmale bei den Milchrinderrassen für die vit mit der Zuchtwertschätzung beauftragt ist. IT-Solutions for Animal Production, Vereinigte Informationssysteme Tierhaltung w.V. (vit), 27283 Verden, Germany, https://www.vit.de/fileadmin/DE/Zuchtwertschaetzung/Zws_Bes_deu.pdf,access: 20.07.-29.09.2022.

Von Engelhardt, W., Breves, G., Diener, M., and Gäbel, G. (2015). Physiologie der Haustiere. 5. Auflage. Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany..

Wallace, R. J., Sasson, G., Garnsworthy, P. C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A. R., Strozzi, F., and Biscarini, F. (2019). A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Science Advances* 5(7):83-91.

Wang, L., Chang, J. H., Paik, S. Y., Tang, Y., Eisner, W., and Spurney, R. F. (2011). Calcineurin (CN) activation promotes apoptosis of glomerular podocytes both in vitro and in vivo. *Molecular Endocrinology* 25(8):1376-1386.

Wang, T., Xue, X., Xie, X., Ye, K., Zhu, X., and Elston, R. C. (2018). Adjustment for covariates using summary statistics of genome-wide association studies. *Genetic Epidemiology* 42(8):812-825.

Wattiaux, M. A. (1998). Protein metabolism in dairy cows. Technical dairy guide nutrition. 2nd. Madison, WI, USA: The Babcock Institute for International Dairy Research and Development, University of Wisconsin, https://kb.wisc.edu/images/group226/52745/5.ProteinMetabolisminDairyCows.pdf.

Weimer, P., Stevenson, D., Mantovani, H., and Man, S. (2010). Host specificity of the ruminal bacterial community in the dairy cow following near-total exchange of ruminal contents. *Journal of Dairy Science* 93(12):5902-5912.

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., and Birmingham, A. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5(1):1-18.

Wood, G., Boettcher, P., Jamrozik, J., Jansen, G., and Kelton, D. (2003). Estimation of genetic parameters for concentrations of milk urea nitrogen. *Journal of Dairy Science* 86(7):2462-2469.

Xue, M.-Y., Sun, H.-Z., Wu, X.-H., Liu, J.-X., and Guan, L. L. (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):1-19.

Yang, S., Luo, J., Chen, Y., Wu, R., Liu, H., Zhou, Z., Xiao, Y., and Shi, D. (2022). A buffalo rumen-derived probiotic (SN-6) could effectively increase Simmental growth performance by regulating fecal microbiota and metabolism. *Research Square*:PPR (preprint) no. 487347.

Zhang, Y., Liu, J., Li, W., Cao, F., Niu, G., Ji, S., Du, X., Cao, B., and An, X. (2020). A Regulatory Circuit Orchestrated by Novel-miR-3880 Modulates Mammary Gland Development. *Frontiers in Cell and Developmental Biology* 8:383.

11 Appendix

11.1 Identification of genomic regions influencing N metabolism and N excretion in lactating Holstein-Friesians (1st study)

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 July 2021
doi: 10.3389/fgene.2021.699550

Identification of Genomic Regions Influencing N-Metabolism and N-Excretion in Lactating Holstein-Friesians

*Hanne Honerlagen[1], Henry Reyer[1], Michael Oster[1], Siriluck Ponsuksili[1], Nares Trakooljul[1], Björn Kuhla[2], Norbert Reinsch[3] and Klaus Wimmers[1,4]**

[1] Genomics Unit, Institute for Genome Biology, Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany, [2] Metabolism Efficiency Unit, Institute of Nutritional Physiology "Oskar Kellner," Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany, [3] Livestock Genetics and Breeding Unit, Institute of Genetics and Biometry, Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany, [4] Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

Excreted nitrogen (N) of dairy cows contribute to environmental eutrophication. The main N-excretory metabolite of dairy cows is urea, which is synthesized as a result of N-metabolization in the liver and is excreted via milk and urine. Genetic variation in milk urea (MU) has been postulated but the complex physiology behind the trait as well as the tremendous diversity of processes regulating the N-metabolism impede the consistent determination of causal regions in the bovine genome. In order to map the genetic determinants affecting N-excretion, MU and eight other N-excretory metabolites in milk and urine were assessed in a genome-wide association study. Therefore phenotypes of 371 Holstein- Friesians were obtained in a trial on a dairy farm under near commercial conditions. Genotype data comprised SNP information of the Bovine 50K MD Genome chip (45,613 SNPs). Significantly associated genomic regions for MU concentration revealed *GJA1* (BTA 9), *RXFP1,* and *FRY1* (both BTA 12) as putative candidates. For milk urea yield (MUY) a promising QTL on BTA 17 including *SH3D19* emerged, whereas *RCAN2*, *CLIC5*, *ENPP4*, and *ENPP5* (BTA 23) are suggested to influence urinary urea concentration. Minor N-fractions in milk (MN) may be regulated by *ELF2* and *SLC7A11* (BTA 17), whilst *ITPR2* and *MYBPC1* (BTA 5), *STIM2* (BTA 6), *SGCD* (BTA 7), *SLC6A2* (BTA 18), *TMCC2* and *MFSD4A* (BTA 16) are suggested to have an impact on various non-urea-N (NUN) fractions excreted via urine. Our results highlight genomic regions and candidate genes for N-excretory metabolites and provide a deeper insight into the predisposed component to regulate the N-metabolism in dairy cows.

Keywords: GWAS, emission, cattle, nitrogen, milk urea, non-urea nitrogen

OPEN ACCESS

Edited by:
Simone Ceccobelli, Marche Polytechnic University, Italy

Reviewed by:
Hadi Atashi, Shiraz University, Iran
Francesco Tiezzi, North Carolina State University, United States

***Correspondence:**
Klaus Wimmers
wimmers@fbn-dummerstorf.de

Specialty section:
This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics

Received: *23 April 2021*
Accepted: *18 June 2021*
Published: *14 July 2021*

Citation:
Honerlagen H, Reyer H, Oster M, Ponsuksili S, Trakooljul N, Kuhla B, Reinsch N and Wimmers K (2021) Identification of Genomic Regions Influencing N-Metabolism and N-Excretion in Lactating Holstein-Friesians. Front. Genet. 12:699550. doi: 10.3389/fgene.2021.699550

INTRODUCTION

Dairy farming contributes with considerable amounts to total environmental nitrogen (N) emissions 2020 (Uwizeye et al., 2020). Due to societal demands and political decisions, the dairy sector is challenged to reduce its N-emissions in the near future. Generally, N-emissions in dairy farming arise after digestion and metabolization of dietary crude protein (CP) (Bryant, 1970; Abdoun et al., 2006). In detail, most of CP is digested by the microbiota

to peptides, amino acids and ammonia (NH3). Depending on the availability of energy sources in the rumen, ammonia is utilized by microbes producing microbial protein (MOP) and facilitating microbial growth. The MOP and rumen undegradable protein are subsequently transported to and digested in the cow's small intestine. The resulting amino acids and smaller peptides are absorbed by the intestinal epithelia. Non-digested dietary N, microbial N originating from the large intestine as well as endogenous N from the cow's metabolism are excreted with feces (Nfec). Surplus ruminal ammonia is transported through the rumen epithelia and by blood proteins to the liver, where it is metabolized to urea by hepatocytes. Urea is excreted by urine and milk, or is transported via the blood stream across the rumen wall or saliva into the rumen. Ruminal urea is instantaneously converted to ammonia resulting in further MOP synthesis, independent of the dietary crude protein level (Munyaneza et al., 2017). Due to the dynamics of N-cycling, the concentrations of blood urea (BU), urinary urea (UU), and milk urea (MU) are directly correlated (Burgos et al., 2007). While quantitative measurements of UU excretions are difficult to carry out in practice, MU values are collected routinely in monthly milk records. Thereby MU is measured via infrared technology, which probably does not provide the highest accuracy but makes measurements available at a large scale. Thus, MU is considered a convenient parameter to assess N-excretion and N-use efficiency in large populations with multiple measurements per individual cow (Burgos et al., 2007, 2010).

In consequence of the cooperative N-usage between the cow and the microbes in the digestive tract, all N-excretions including MU, UU and Nfaec mean a loss of bioavailable N (Munyaneza et al., 2017). Besides urea, further N-fractions are excreted via urine and milk, comprising purine derivatives, creatine and creatinine, as well as hippuric acid and ammonia (Cox, 2013; Munyaneza et al., 2017). Although these non-urea-nitrogen (NUN) metabolites account only for a small part of total N-excretion, they also contribute to N-losses and N-emissions into the environment (Jardstedt et al., 2017). Generally, there is high variation in NUN-excretion (Cox, 2013), but also evidence for phenotypic variation in MU concentration, which is most commonly explained by the feed ration. Interestingly, Aguilar et al. (2012) referred to phenotypic variation in MU concentrations that could not be explained by ration composition and determined the cow itself as a highly significant factor influencing MU concentrations. The feed-independent variation in MU concentrations has led to the assumption that genetic components might have a significant impact on the N-metabolism and thus on N-excretion of dairy cows.

Indeed, several studies have calculated moderate heritabilities for MU concentrations in Holstein cows, ranging from 0.13 to 0.59 (Wood et al., 2003; Stoop et al., 2007; Stamer et al., 2011; Guarini et al., 2019). Average heritability estimates over a whole lactation comprised values from 0.24, 0.22 to 0.2 for Holsteins in first, second and third lactation, respectively (Jahnel et al., 2021). Estimations of the genetic relationship between MU concentrations and other milk production traits, such as milk-, fat-, and protein yield (Wood et al., 2003; Stamer et al., 2011; Höfener, 2019), and fat and protein percentage (Höfener, 2019) showed only little or even no correlation. Considering these facts, MU concentrations have a potential for breeding intervention based on genotypes that might determine a more efficient N-metabolism. Interestingly, a conventional breeding value for MU concentrations was calculated in consequence of political decisions in the Netherlands (Sebek et al., 2007) but not much has been published yet on the underlying genome regions. Only in a few studies, quantitative trait loci (QTL) for MU have been detected. For example, Bouwman et al. (2010) localized four genomic regions on BTA 1, 6, 21 and 23 for MU concentration and MU yield, using genotypes and test day data from 1926 Holstein-Friesians in the Netherlands. Cecchinato et al. (2014) detected three significantly associated SNPs on BTA 1, 7, and 26 in Brown Swiss cattle, while Pegolo et al. (2017) found regions on BTA 4, 5, and 13 significantly related to MU concentrations.

Apparently, the functional biodiversity and the quantitative character of MU make it difficult to determine genomic regions consistently associated with N-metabolism in dairy cows. Consequently, in the current study, in addition to the commonly investigated traits (MU concentration and MU yield), further N-excretion metabolites that are involved in N-metabolism were investigated in a genome-wide association approach. Specifically, the list of traits comprised MU, MU yield (MUY), urinary urea (UU), total minor N-metabolites in milk (MN) and four specific NUN-fractions in urine. The aim of our study was to uncover genetic variants and genomic regions that are significantly associated with MU and other N-excretion and N-metabolism traits in order to better understand the genetic background of N-metabolism in dairy cows.

MATERIALS AND METHODS

Animal husbandry and sampling were carried out according to the guidelines of the German Animal Protection Law. All protocols were approved by the Institute's Animal Welfare Commission. The experimental protocol is in strict compliance with the German Animal Welfare Legislation, has been approved by the Ethics Committee of the federal state of Mecklenburg-Western Pomerania, Germany (Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei; LALLF M-V7221.3-2-019/19) and is in accordance with the ARRIVE guidelines.

Cow Population

For this study, all 371 lactating Holstein-Friesians owned by the practice operating dairy farm Gut Dummerstorf (Mecklenburg-Western Pomerania, Germany) were investigated. The entire herd was housed in a free stall pen, fed twice daily (6.00 a.m./p.m., *ad libitum*) and milked twice daily (4.00 a.m./p.m.) in a 2 × 14 side-by-side milking parlor. All cows were fed the same ration, which provided 7.26 MJ NEL/kg dry matter and 15.4% crude protein resulting on average in a slightly negative ruminal nitrogen balance of –0.9 g N/kg dry matter. The composition of the ration was constant throughout the trial. The cows were in 1st to 9th lactation and at least 21 days in milk. The average herd's daily milk yield was 36.2 kg per day with an average

lactation duration of 364 days. The cows descended from 145 sires and 163 dam sires.

Sample Collection and Analysis

The trial was carried out in four subsequent sampling time points in spring 2020. Milk samples were obtained as pooled samples of morning (25 ml) and evening milking (25 ml) for each cow in the monthly milk record procedure (in total 50 ml per cow). Individual urine samples (10 ml) were collected in a maximum time distance of 24 h relative to the monthly milk record procedure (n = 371). The cows' body weights were recorded two times during the entire trial right after the morning milking.

Milk constituents were measured by the State Control Federation of Mecklenburg-Western Pomerania. From each cow, the pooled milk sample (50 ml) was conserved in a Bronopol/Kathon mixture and stored at 4°C for not longer than 48 h before urea content was determined using mid infrared spectroscopy (CombiFoss 7, Foss, Hilleroed, Denmark). Therefore, milk samples were homogenized and heated to 40°C and the absorption measured at the corresponding wavelength was electronically transformed on the basis of a regression calculation (Partial Least Squares) into MU concentrations (mg/l). Furthermore, data including protein, fat, lactose and milk yield of monthly milk records were available from the herd.

In addition, the total N content of milk was analyzed by MQD (Qualitätsprüfungs- und Dienstleistungsgesellschaft Mecklenburg-Vorpommern, Güstrow, Germany) according to the Kjeldahl method in accordance with ASU §64 LFGB L 01.00–10/1. Results were converted to milk crude protein (g/100 g), using the factor 6.38.

Urine samples were obtained right after morning or evening milking by dorso-ventral stimulating massages, ventral of the vulva and during spontaneous urinating. The urine was immediately stored in 15 ml tubes on ice, before it was frozen at –20°C. Frozen urine samples were thawed on ice, homogenized and centrifuged at 13,000 × g and 4°C for 5 min before they were analyzed for urea and NUN-metabolites. Urea was diluted 20-fold in water and analyzed photometrically at a ABX Pentra C400 clinical chemistry analyzer (HORIBA Europe GmbH, Oberursel, Germany). The NUN-metabolites were measured by HPLC (1200/1260 infinity Series; Agilent Technologies, Waldbronn, Germany) as previously described (Müller et al., 2021). Detection and quantification of NUN-metabolites were conducted at 230 nm with exception for allantoin and creatine at 210 nm.

Phenotypic Traits

Nine phenotypes, divided into repeatedly measured traits (RMTs) and single measured traits (SMTs), were generated from milk and urine data. RMTs represent the entire cow individual lactation by using the available milk record data of 14 months. SMTs are derived from one-time measurements during the sampling trial. RMTs using milk data are MU in mg/l and MUY in g. For MU, values of 14 months milk record data were averaged for each cow. MUY was generated by multiplying the individual daily milk yield with the MU value of each test day in the 14 months period, summed up and averaged. Milk nitrogen (MN) is the only single-measured milk trait and contains all N-fractions in the milk, i.e., protein-N and non-protein-N (NPN) in g/100 g milk.

All urine traits represent SMTs, expressed in mmol/l urine. Urinary traits comprise UU and the purine derivatives allantoin (AL) and its precursor uric acid (UA). Furthermore, hippuric acid (HA) and creatinine (CRE) as well as its precursor creatine (CR) were quantified.

Genotype Data

The whole cow population of Gut Dummerstorf was genotyped in the course of the German-wide genotyping project KuhVision (Rensing et al., 2017). SNP chip data of the project was imputed by vit Verden (Vereinigte Informationssysteme Tierhaltung, Verden, Germany) on the Bovine 50K MD Genome chip (45,613 SNPs). The imputing was conducted as part of the genetic evaluation routine with 825,999 Holstein cows and pedigree information of 2,467,138 animals, using FIMPUTE version 3.0 (Sargolzaei et al., 2010). The imputing accuracy per chromosome was reported to be 99.5–99.7% by vit Verden. Probe sequences of the SNP chip were mapped to the current bovine genome assembly and the annotation of markers was updated to ARS-UCD1.2 (access 01 September 2020). Markers not mapping to autosomes in ARS-UCD1.2 were discarded (1,940 SNPs). Furthermore, genotype data was filtered for minor allele frequency (MAF) > 0.05 and deviation from Hardy Weinberg equilibrium ($p > 1 \times 10^{-6}$) (4,220 SNPs discarded). After imputing, mapping and filtering, 39,453 SNPs remained for genome-wide association analysis.

Genome-Wide Association Study (GWAS)

A two-step approach was utilized to identify genomic regions associated with nitrogen phenotypes in milk and urine conducting the GAPIT R package with the BLINK method and default parameters (Lipka et al., 2012; Huang et al., 2019). In a first step, all phenotypes were adjusted for environmental effects in a linear model, giving residuals for each trait. The phenotypes were adjusted as follows: For MU and MUY, the average milk fat content of the lactation was included as covariate in the model accounting for differences in feed intake between individuals, considering the slightly negative RNB in the feed ration. For MU, lactation number, which separates cows in their first lactation from other cows, was considered as fixed effect in the model to correct for physiological protein anabolism in the course of growth during first lactation. Lactation number was not included in the model for MUY, due to a confounding with milk yield, which is a component of MUY calculation. For MN, milk protein and MU content of the test day were integrated as covariates in the model to investigate only the NPN-and NUN-fractions in milk. In addition, for MN and for UU, the milk fat content of the test day was included as covariate to account for variation in feed intake. For all urinary traits, except for CRE, the covariates CRE and body weight were used to correct for differences in urine volume. CRE was adjusted for deviation in muscle metabolism with body weight as covariate.

In a second step, GWAS was performed with the trait residuals by using the Bayesian information and Linkage-disequilibrium

Iteratively Nested Keyway (BLINK) method. BLINK executes two fixed effect models in an iterative approach. One of which accounts for population stratification by testing each SNP one by one with multiple associated SNPs that are fitted as covariates on the testing SNP. The second model selects the covariate SNPs instead of a kinship matrix to correct on the relationship in the population investigated. With this approach, BLINK abolishes the necessity for the genes underlying a trait to be evenly distributed along the genome in order to enhance statistical power (Huang et al., 2019). For each trait, a quantile-quantile (QQ) plot was generated, based on the observed *P*-values.

To account for linkage disequilibrium between markers, the R script SimpleM was used to estimate the number of independent tests. SimpleM calculates a composite linkage disequilibrium matrix from the SNP genotypes and uses this matrix in a principal component analysis to determine the effective number of independent tests (Gao et al., 2008). For the data set 20,517 independent tests were ascertained. The significance thresholds were consequently set at 1/20,517 [–log10 (*P*-value) = 4.31] for suggestive significance and at 0.05/20,517 [–log10 (*P*-value) = 5.61] for genome-wide significance (Lander and Kruglyak, 1995). Genome-wide results for each trait were visualized in Manhattan plots.

Data Integration and Candidate Genes

Based on the results from GWAS, the genomic regions surrounding the significant SNPs were examined for each trait. In accordance with Saatchi et al. (2014), 1 Mb regions were defined around the significant SNPs (500 kb each direction from the SNP) as putative QTL regions. The QTLs were manually investigated for the two genes closest to the right and left of the significant SNPs and for positional (covering significantly associated SNP) and functional (postulated function with relevance to the trait) candidate genes. Genomic information and functional annotations of genes were extracted from GeneCards[1] and Ensembl[2].

[1]http://www.genecards.org

[2]http://www.ensembl.org

RESULTS

Both the milk and urine traits showed appreciable variation in the Holstein population studied (**Table 1**). However, urine traits showed a considerably higher variability compared to milk traits. This might be due to varying urine volumes, which led to a dilution or concentration of the urine components (Spek et al., 2012). QQ plots from GWAS are shown in **Supplementary Figure 1** for all traits, indicating that the approach used was suitable to control for population stratification.

Urea in Milk and Urine

For MU, three significant SNPs were identified, which mapped on BTA 9 (rs41609177) and on BTA 12 (rs110792428, rs110799535) (**Table 2** and **Figure 1A**). On BTA 12 both significant SNPs are located in the same QTL region between 28.8 and 29.9 Mb, which comprises six putative candidate genes. The region on BTA 9 indicated at 29.8 Mb contains the annotated genes gap junction protein alpha 1 (*GJA1*) and domain family member 32 (*TBC1D32*).

For MUY as the consolidating trait of milk urea and milk yield, rs41835093 and rs41835125 were found to be significantly associated (**Table 2** and **Figure 1B**). The two SNPs mapped on BTA 17 at 6.6 Mb, only 0.03 Mb apart from each other. They are located in the positional candidate gene SH3 domain containing 19 (*SH3D19*). Furthermore, in close vicinity six other genes are annotated in this QTL region.

For UU, as the quantitatively most relevant nitrogen fraction in urine, rs109244808 on BTA 23 indicated a single QTL at 18.9 Mb (**Table 2** and **Figure 1C**). The QTL harbors five genes, including chloride intracellular channel 5 (*CLIC5*) and regulator of calcineurin 2 (*RCAN2*).

Non-urea-N (NUN) in Milk and Urine

For MN, which represents minor NPN and NUN-metabolites in milk, QTLs on BTA 4, 16, and 17 were indicated by five significantly associated SNPs (**Table 2** and **Figure 2A**).

TABLE 1 | Descriptive statistics of milk and urinary nitrogen traits in Holstein cows.

Acronym	Trait	Unit	Matrix	Analysis	Min	Max	Mean	SD	N
MU	Milk urea	mg/l	Milk	RMT[a]	121.00	274.54	192.53	24.80	371
MUY	Milk urea yield	g	Milk	RMT	3.64	10.53	6.84	1.37	371
UU	Urinary urea	mmol/l	Urine	SMT	13.40	308.00	102.54	49.63	331[c]
MN	Milk nitrogen	g/100 g	Milk	SMT[b]	2.70	4.75	3.62	0.38	368[d]
UA	Urine acid	mmol/l	Urine	SMT	0.40	3.06	1.52	0.31	331[c]
AL	Allantoin	mmol/l	Urine	SMT	1.59	31.68	18.01	5.48	331[c]
HA	Hippuric acid	mmol/l	Urine	SMT	5.99	60.70	27.04	9.93	331[c]
CR	Creatine	mmol/l	Urine	SMT	0.49	12.61	3.64	1.80	331[c]
CRE	Creatinine	mmol/l	Urine	SMT	0.75	9.35	4.35	1.59	331[c]

[a]*RMT, repeated measurement trait.*

[b]*SMT, single measurement trait.*

[c]*Reliable body weight measurements could only be recorded in 331 out of 371 cows, resulting in 331 phenotypes for urine traits since body weight was included as covariate in the urine models.*

[d]*Three samples out of initially 371 had been discarded in laboratory analysis resulting in 368 MN phenotypes*

TABLE 2 | Overview of QTLs and suggested candidate genes for N-traits identified by GWAS.

Trait	BTA[a]	QTL interval (Mb)[b]	Significant SNP	$-\log_{10}$ (*P*-value)	MAF[c]	R^2[d]	Candidate gene[e]
MU	9	29.29–30.29	rs41609177	4.73	0.26	0.06	*GJA1, TBC1D32*
MU	12	28.84–29.89	rs110792428	4.50	0.42	0.06	*RXFP2, FRY*
			rs110799535	4.86	0.41	0.06	
MUY	17	6.12–7.15	rs41835093	4.81	0.31	0.07	*SH3D19, PRSS48, LRBA, RPS3A*
			rs41835125	4.36	0.38	0.06	
UU	23	18.49–19.49	rs109244808	4.54	0.14	0.05	*CLIC5, RCAN2, ENPP4, ENPP5*
MN	4	20.65–21.65	rs43375405	4.91	0.44	0.05	
MN	16	57.55–58.85	rs41814676	4.52	0.10	0.04	*ASTN1*
			rs41814682	4.52	0.10	0.04	*PAPPA2*
MN	17	18.57–19.57	rs41576712	5.89	0.188	0.06	*SLC7A11*
MN	17	17.56–18.56	rs109749663	4.85	0.06	0.05	*MGST2, ELF2*
UA	7	63.30–64.30	rs42372803	5.45	0.36	0.06	*NMUR2*
HA	7	38.87–39.87	rs109848970	8.82	0.07	0.09	
HA	12	74.36–75.36	rs43706910	5.96	0.32	0.05	
HA	17	69.78–70.78	rs41634411	5.59	0.37	0.06	*SFI1*
CR	4	116.16–117.16	rs109873598	4.37	0.24	0.05	*DPP6*
CR	5	82.0–83.00	rs110770413	4.85	0.31	0.06	*ARNTL2, ITPR2*
CR	5	65.16–66.16	rs41590238	4.83	0.06	0.06	*MYBPC1*
CR	6	46.21–47.21	rs43705592	4.31	0.32	0.06	*STIM2*
CR	7	63.44–64.44	rs110536156	5.12	0.40	0.06	
CR	7	67.16–68.16	rs42332347	4.53	0.39	0.05	*SGCD*
CR	18	23.62–24.62	rs109756131	4.74	0.41	0.06	*SLC6A2*
CRE	16	2.51–3.51	rs41787830	4.35	0.15	0.06	*TMCC2, MFSD4A*

[a] *Bos taurus autosome.*
[b] *Compiled according to Saatchi et al. (2014).*
[c] *Minor allele frequency of the significant SNP.*
[d] *Calculated by linear regression between phenotype residuals and SNP-alleles.*
[e] *Selected due to a positional or functional association to the trait; MU, milk urea concentration; MUY, milk urea yield; UU, urinary urea; MN, minor N-fractions (non-protein, non-urea) in milk; UA, uric acid; HA, hippuric acid; C, creatine; CRE, creatinine.*

One of them, rs41576712 on BTA 17, exceeded genome-wide significance. The same chromosome harbors a further QTL region, with rs109749663 at 18.06 Mb highlighted as significantly associated. This SNP indicates microsomal glutathione S-transferase 2 (*MGST2*) as positional candidate. On BTA 16, rs41814676 and rs41814682 mapped only 0.03 Mb apart from each other in astrotactin 1 (*ASTN1*) and pappalysin 2 (*PAPPA2*) at 58 Mb. The significantly associated SNP on BTA 4 at 21.15 Mb points on ADP ribosylation factor like GTPase 4A (*ARL4A*) and scinderin (*SCIN*) as putative candidate genes in the vicinity.

For AL no SNP exceeded the significance threshold, but for UA one SNP occurred significant (**Table 2** and **Figures 2B,C**). This SNP (rs42372803) is located on BTA 7 at position 63.8 Mb and represents an intron variant of the long non-coding RNA ENSBTAG000054682. However, in close proximity, neuromedin U receptor 2 (*NMUR2*) provides a putative functional candidate gene.

For HA, as a by-product of phenolic acid digestion excreted by urine, the analysis revealed two SNPs that reached the genome-wide significance threshold. These are rs109848970 (P = 8.82) on BTA 7 and rs43706910 (P = 5.96) on BTA 12 (**Table 2** and **Figure 3A**). Although neither of these two SNPs directly mapped in a gene region, rs109848970 opens up a QTL at 39.4 Mb (BTA 7) with 15 annotated genes. In the immediate vicinity of rs43706910 at 74.86 Mb on BTA 12 six genes are annotated so far. A third SNP on BTA 17 (rs41634411) was significantly associated with HA. This SNP is located at 70.28 Mb in the gene encoding for the SFI1 centrin binding protein (*SFI1*).

For CR, which plays an indispensable role in muscular energy metabolism, the urinary levels in the cow herd were significantly associated with seven SNPs located on BTA 4, 5, 6, 7, and 18 (**Table 2** and **Figure 3B**). Two QTLs on BTA 5 and one QTL on BTA 18 revealed more than 10 annotated genes, respectively, while the other QTL regions included less than five annotated genes, each. Three significantly associated SNPs, that are rs110770413 (BTA 5 at 82.5 Mb), rs42332347 (BTA 7 at 67.7 Mb) and rs109873598 (BTA 4 at 116.7 Mb) mapped in gene regions and indicated aryl hydrocarbon receptor nuclear translocator like 2 (*ARNTL2*, BTA 5), sarcoglycan delta (*SGCD*, BTA 7) and dipeptidyl peptidase like 6 (*DPP6*, BTA 4) as positional candidates.

For CRE, which is an excretory derivative of CR, one SNP (rs41787830) on BTA 16 reached the threshold of suggestive significance (**Table 2** and **Figure 3C**). This SNP is located at 3.0 Mb in an intron of the gene coding for transmembrane and coiled-coil domain family 2 (*TMCC2*). Furthermore, this QTL includes another 16 annotated genes.

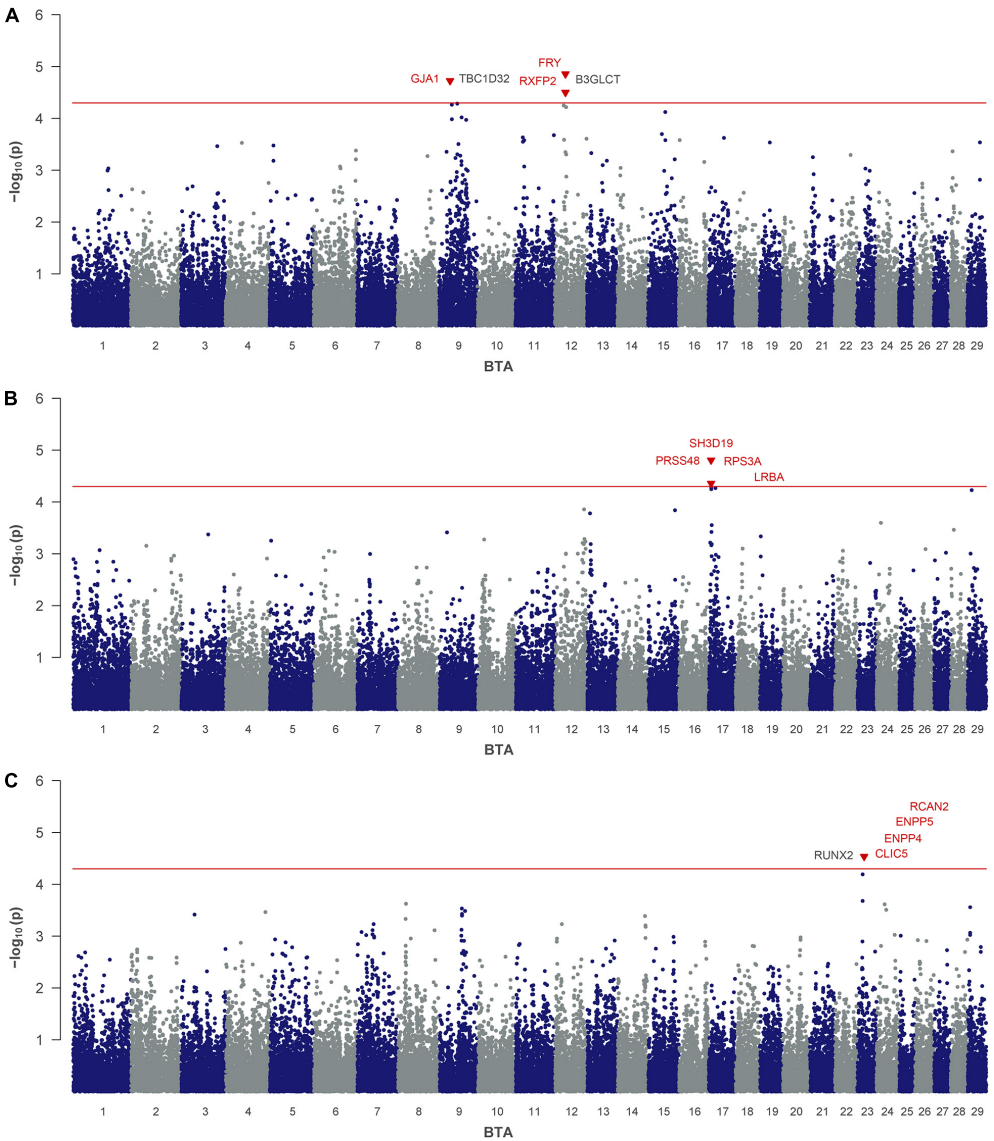

FIGURE 1 | Genome-wide association analysis of urea traits: **(A)** milk urea, **(B)** milk urea yield and **(C)** urinary urea. The red line indicates the threshold for suggestive significance. Genes with a suggested positional or biological association to the trait are indicated in red. Genes, located closest to the right and left of the lead SNP in a 1-Mb window are colored in black.

DISCUSSION

As a part of the effort to reduce N-emissions in dairy cows, this work analyzed the genetics of several relevant traits that contribute to the variance in N-excretion. For this purpose, a herd of 371 Holstein cows was comprehensively phenotyped for nine N-fractions from milk and urine and investigated in a GWAS approach. Significantly associated genomic regions and potential candidate genes were identified for eight of these fractions.

Urea in Milk and Urine

N-excretion is generally influenced by the cow and its rumen microbiome. The most relevant N- excretion fraction in terms of quantity is urea (Munyaneza et al., 2017). Besides the microbiological impacts in the rumen, mainly the following processes are conceivable to influence the amount of excreted and recycled nitrogen in the cow: (i) absorption of ammonia via the rumen epithelia, transport to the liver by blood proteins and hepatic synthesis of urea; (ii) transport of synthetized urea from liver via blood (BU) to kidney, udder and rumen; (iii) diffusion and excretion of BU via nephrons (UU) and the udder epithelium (MU).

For MU, significant genomic regions were identified on BTA 9 and 12. The region on BTA 9 indicates *GJA1* as a potential candidate. *GJA1* encodes connexin 43, a component of gap junctions that contribute to the diffusion of low molecular weight substances into cells (Graham and Simmons, 2005). Therefore, it seems conceivable that *GJA1* also influences the diffusion of

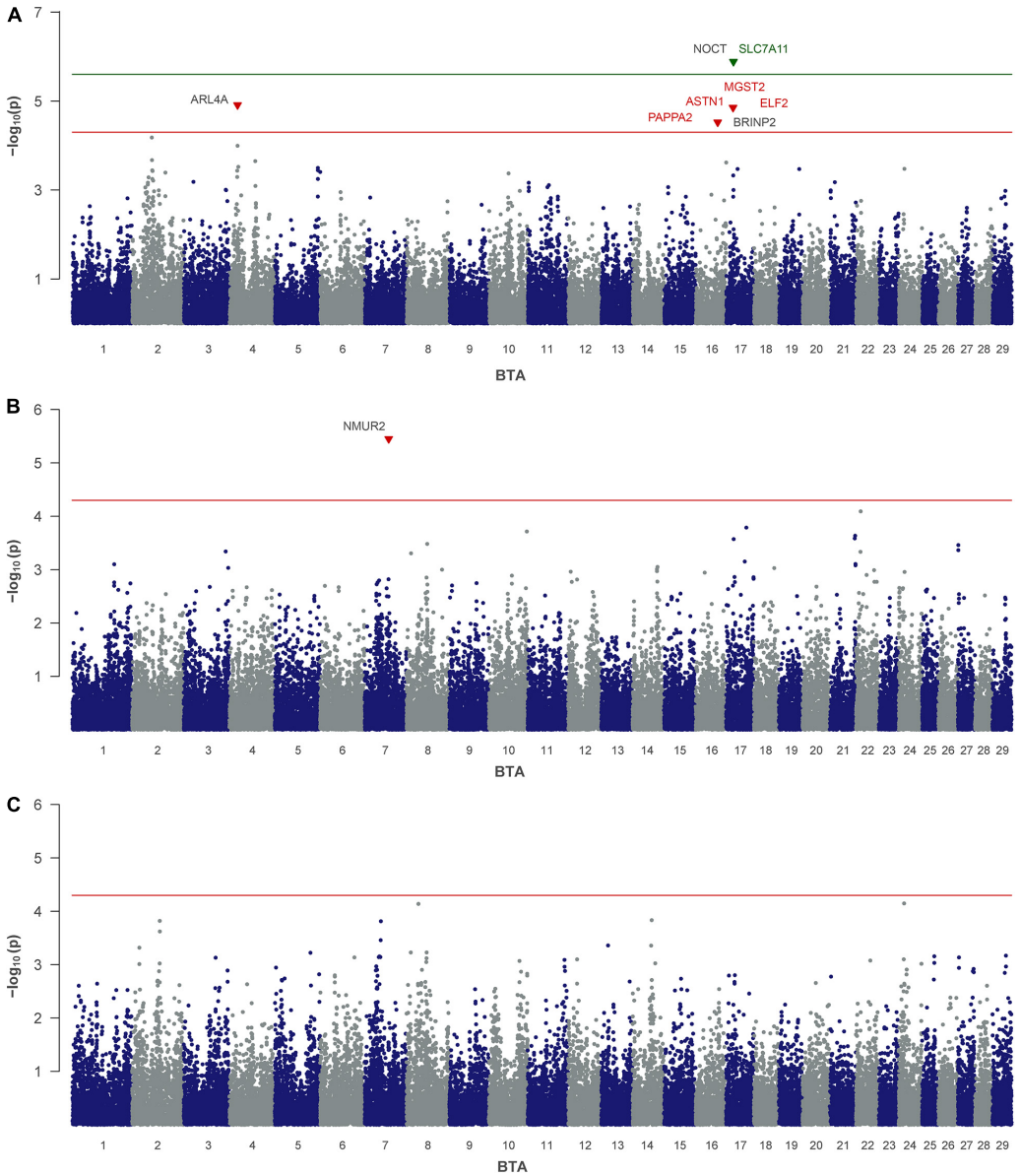

FIGURE 2 | Genome-wide association analysis of non-urea-nitrogen traits: **(A)** milk nitrogen, **(B)** uric acid and **(C)** allantoin. The red and green lines indicate suggestive and genome-wide significance thresholds, respectively. Genes with a suggested positional or biological association to the trait are indicated in red and green. Genes located closest to the right and left of the lead SNP in a 1-Mb window are colored in black.

urea across the udder epithelium, and thus the urea content of the milk. Interestingly, connexin 43 was previously shown to play a critical role in the development of the mammary gland epithelia and in lactation processes in mice and cattle (Plante and Laird, 2008; Farhadian et al., 2021). Moreover, *GJA1* has been associated with fertility parameters in cows (Ribeiro et al., 2016; Neupane et al., 2017). In the second QTL identified for MU on BTA 12, a number of candidate genes related to cellular turnover processes are mapped and annotated. Interestingly, the udder epithelia of cows is subject to strong cellular turnover events during lactation. Specifically, in early lactation massive cell proliferation causes the increase in milk yield, while apoptosis in late lactation strongly influences the dry period induction (Stefanon et al., 2002; Hale et al., 2003). Considering that cellular turnover processes have an impact on the diffusion capacities in the udder epithelium, an influence of those genes on MU levels seems to be conceivable. In this context relaxin family peptide receptor 2 (*RXFP2*) was shown to be significantly associated with mammary gland morphology in Simmental cattle (Pausch et al., 2016). Moreover, FRY microtubule binding protein (*FRY*) has been confirmed to influence the development and function of the mammary gland (Liu et al., 2019).

Milk urea yield was calculated as the absolute MU volume over the entire lactation period, i.e., even with a lower MU per

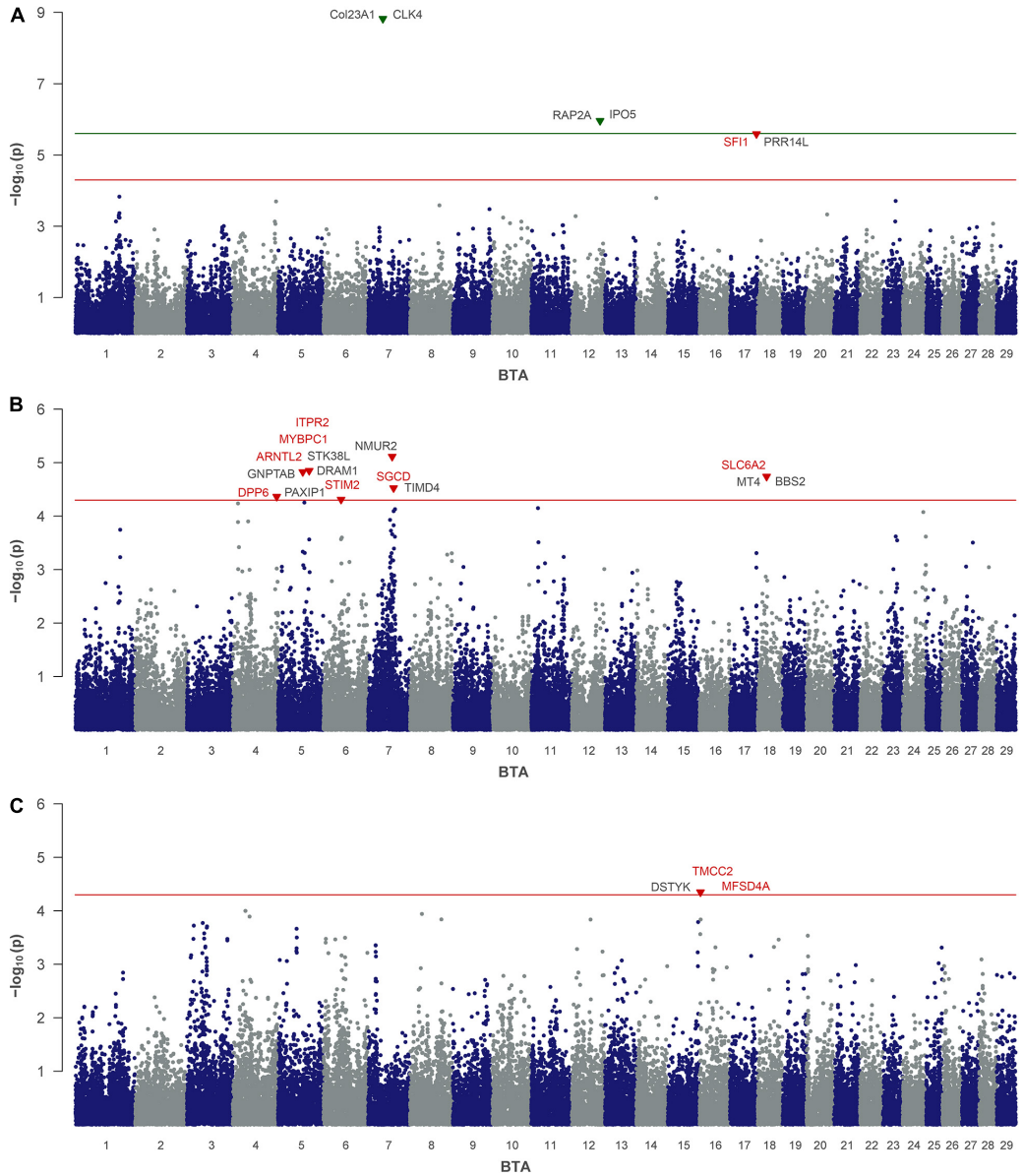

FIGURE 3 | Genome-wide association analysis of non-urea-nitrogen traits: **(A)** hippuric acid, **(B)** creatine and **(C)** creatinine. The red and green lines indicate suggestive and genome-wide significance thresholds, respectively. Genes with a suggested positional or biological association to the trait are indicated in red. Genes, located closest to the right and left of the lead SNP in a 1-Mb window are colored in black.

liter, a cow with a high milk yield can have a comparatively high MUY. Two significant SNPs on BTA 17 revealed *SH3D19* as positional candidate for MUY. Although *SH3D19* has not been postulated in the context of urea or milk yield so far, it has been mentioned in several studies concerning fertility traits in cattle and pigs (Kranc et al., 2017; Guarini et al., 2019; Brazert et al., 2020). Fertility traits and milk yield basically compete for available energy in early lactation of high-yielding cows. Genomic correlations between milk yield and fertility parameters as well as milk urea and fertility traits have already been proven by several studies (König et al., 2008; Mucha and Strandberg, 2011; Sawa et al., 2011). It is also worth noting that other genes in this QTL such as serine protease 48 (*PRSS48*), LPS responsive beige-like anchor protein (*LRBA*) and ribosomal protein S3A (*RPS3A*) have also been mentioned in the context of fertility. In fact, *RPS3A* serves as an indispensable factor in reproductive processes not only in cattle but also in sheep (He et al., 2017; Pokharel et al., 2020). It is therefore conceivable that not only the positional candidate *SH3D19* but the entire QTL has an influence on fertility, milk urea and milk yield (Hu et al., 2019). More detailed investigations of this genomic region toward the influence on energy metabolism as the common denominator of MUY and fertility traits are proposed.

For UU a single QTL on BTA 23 revealed six candidate genes. Four of which (*RCAN2*, *CLIC5*, *ENPP4*, and *ENPP5*) are of relevance due to their biological functions in the nephron. Filtration and secretion processes in the nephron determine the amount of excreted UU. Both, the filtration of urea in the glomerus, the first section of the nephron, and the reabsorption of urea back into the blood in the last section of the tubule, depend on concentration-dependent diffusion. *RCAN2*, which regulates the release of calcineurin, may be an obvious candidate for UU because calcineurin activates podocyte (glomerus cell) apoptosis (Wang et al., 2011). This might affect the absolute renal diffusion rate for urea in the glomerus, resulting in an altered amount of urea in urine. *CLIC5* is indispensable for the development and preservation of glomerular endothelial cells and podocytes. In addition, chloride transporters in general play an important role in establishing the concentration gradient between blood and urine, which is relevant for filtration and reabsorption of urea (Schwiebert et al., 1994; Jentsch et al., 2002). Finally, ectonucleotide pyrophosphatase phosphodiesterase 4 and 5 (*ENPP4* and *ENPP5*) might play a key role in the activation of purinoceptors, which are known as renal tubular transporters regulating sodium and water balance (Vekaria et al., 2006). Sodium influences the amount of reabsorbed urea by its osmotic potential, while the water balance is influenced by the amount of reabsorbed urea.

Non-urea-N (NUN) in Milk and Urine

Apart from urea, several other N-metabolites are excreted via milk and urine. These NUN are mainly intermediates and end products of protein and nucleic acid degradation and are known to be highly variable between individuals (Bristow et al., 1992).

Milk nitrogen generally comprises all N-fractions in milk including proteins, urea and NUN. By calculating the residuals considering the effects of MU and milk protein in the model, the present GWAS for MN focuses on the NUN-fractions in milk. A major subset of NUN in milk are free amino acids (Roginski et al., 2003; Rafiq et al., 2016). Among the four QTLs identified for MN, three significantly associated SNPs mapped in *ASTN1*, *PAPPA2,* and *MGST2*. According to current annotations, these genes are involved in neuroblast development, energy metabolism and glutathione metabolism. While *MGST2* has been associated to milk fat metabolism in dairy cows (Gebreyesus et al., 2019), a link to variations in NUN-fractions in milk has not been postulated for any of these three genes so far. However, in the same QTL region, E74 like ETS transcription factor 2 (*ELF2*) seems to be more likely to explain individual variation in MN. *ELF2* has been determined as a major transcription factor driving the formation of milk protein in high yielding dairy cows (Maas et al., 1997). Specifically, *ELF2* mediates the translation of mRNA to protein and thus limiting the rate of milk protein synthesis. Milk protein is generally synthesized by the lactocytes utilizing the influx of amino acids. Unused amino acids in this process are excreted as part of MN. Thus, increased or decreased milk protein synthesis induced by *ELF2* could have a major impact on MN. The bioavailability of limiting amino acids is a further determinant of milk protein synthesis. Solute carrier family 7 member 11 (*SLC7A11*) on BTA 17 encodes for a highly specific amino acid transport system that enables the influx of extracellular cysteine in exchange for intracellular glutamate (Verrey et al., 2004). An impact of *SLC7A11* on amino acid availability might influence milk protein synthesis and thus determining the amount of unused amino acids excreted via MN. In the QTL on BTA 4, neither *ARL4A* nor *SCIN* have been associated with minor N-fractions so far, but *ARL4A* has been postulated as a candidate gene for milk yield, recently (Vijayakumar et al., 2019).

Uric acid (UA) is a subsequent product of purine nucleotide degradation. Precisely, the purine nucleotides are degraded to xanthine and oxidized via xanthine oxidoreductase to UA mainly in the intestinal mucosa and in the liver of cattle (Gonzalez-Ronquillo et al., 2003). Although in ruminants, the purine influx mainly derives from intestinal digestion and absorption of rumen microbial nucleic acids (Balcells et al., 1991) UA is excreted constantly, independent of the microbial protein synthesis and the feeding ration (Susmel et al., 1995). One QTL on BTA 7 has been significantly associated with UA, revealing *NMUR2* in the vicinity. NMU, the encoded neuropeptide receptor, has been attributed a central role in the regulation of feeding behavior, feed intake and body weight in animal models (Howard et al., 2000; Sampson et al., 2018), but an association to UA synthesis or excretion in cattle has not been postulated so far.

No genomic region has been significantly associated with allantoin (AL) excretion in urine. AL is a subsequent product of oxidized UA and the quantitatively highest excretion metabolite of purine bases in cows' urine (Bristow et al., 1992; Cox, 2013). It is known that variance in AL is primarily influenced by the rate of microbial protein synthesis (Kehraus et al., 2006; Tas and Susenbeth, 2007).

Hippuric acid (HA) results from the hepatic conjugation of glycine and benzoic acid. The latter is a product of phenolic acid fermentation in the rumen (Sun et al., 2020). Interestingly, HA content in manure could have a significant impact on N-losses from the soil, making it a valuable trait for further genetic improvement, in addition to nutritional measures to reduce N-emissions (Gardiner et al., 2016). The QTL on BTA 7 contains several genes with functional annotations in extracellular matrix formation and degradation (*Col23A1*, *B4GALT7*, *PHYKPL*), in nuclear ribosomal processes (*HNRNPAB*, *NHP2*, *RMND5B*) and in intracellular transports to the Golgi complex (*TMED9*). The significantly associated SNP in the QTL on BTA 12 mapped between RAP2A, member of RAS oncogene family (*RAP2A*) and importin 5 (*IPO5*). The former gene links intestinal microvilli formation and cell polarization, whereas *IPO5* enables the transport of proteins into the nucleus (Gloerich et al., 2012). Based on the current annotation, the proposed genes provide no direct link to influence the variability in HA. However, the rumen microbiota might have an impact on HA differences by synthesizing divergent benzoic acid levels, but if the significantly associated genes influence the ruminal microbial composition and thus affect benzoic acid and HA levels asks for further research.

Creatine (CR) is synthesized in the liver and acts as an energy storage for nucleoside phosphates in the muscle. Since

the muscular energy metabolism is regulated by a dynamic variety of biological processes, a polygenic influence on CR levels seems conceivable. *SGCD* on BTA 7 and myosin binding protein C1 (*MYBPC1*) on BTA 5 are involved in muscle anabolism and muscle contraction. The protein encoded by *MYBPC1* influences the skeletal muscle contractility by targeting muscle-type creatine kinase to myosin filaments. Furthermore, inositol 1,4,5-trisphosphate receptor type 2 (*ITPR2*) on BTA 5 and stromal interaction molecule 2 (*STIM2*) on BTA 6 play roles in Ca^{2+} muscle storage and release. Creatine phosphate levels are known to influence Ca^{2+} entries in skeletal muscle fibers (Fryer et al., 1995). *ITPR2* encodes a receptor that mediates the mobilization of intracellular Ca^{2+} stores, while STIM proteins act as Ca^{2+} sensors and influence the Ca^{2+} influx. In addition, the solute carrier family 6 member 2 (*SLC6A2*) gene localized in the QTL on BTA 18 is of interest due to the identification of a causal mutation for creatine transporter deficiency in humans (Schiaffino et al., 2005).

Creatinine (CRE) is the hydrolyzed derivative of CR and the main excretion product of creatine in urine. A single SNP on BTA 16 had been significantly associated with CRE, indicating *TMCC2* as positional candidate. *TMCC2* forms complexes with the apolipoprotein E and thus influences proteolytic pathways (Hopkins et al., 2011). Given the fact that CR influences muscle turnover processes, a link to protein degradation is conceivable. Furthermore, major facilitator superfamily domain containing 4A (*MFSD4A*) is involved in the opening of glucose dependent sodium channels in the medulla of the kidney and thus influencing the osmotic gradient between tubule and interstitium. Urea reabsorption in the kidney is dependent on both this osmotic gradient and the opening of sodium channels. Interestingly, higher urinary CRE concentrations have been observed in cows with lower urine and milk urea concentrations (Müller et al., 2021). An interplay of urea and CRE excretion via *MFSD4A*, in the context of total N-excretion might therefore be further explored.

CONCLUSION

This study proposes a number of genomic regions and candidate genes associated with MU concentration and a further eight N-excretory traits in milk and urine of Holstein-Friesians. The results complement the knowledge of genetic determinants regulating the N-metabolism of dairy cows, which influences in its entirety N-emissions in dairy farming. In the light of a sustainable food production the results can be of relevance for a targeted exploitation of predisposed physiological plasticity in N-metabolism of dairy cows.

DATA AVAILABILITY STATEMENT

The data was provided by the German Evaluation Center (VIT, Verden) and is subject to the following licenses/restrictions: the data is proprietary and cannot be released publicly. Requests to access these datasets should be directed to the authors.

ETHICS STATEMENT

The animal study was reviewed and approved by the Ethics Committee of the Federal State of Mecklenburg-Western Pomerania, Germany (Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei; LALLF M-V7221.3–2–019/19.

AUTHOR CONTRIBUTIONS

HH organized the sampling trial. HH and HR carried out the sampling trial, performed statistical analysis and interpretation of the data, and wrote the manuscript. NT, MO, and SP supported data analysis and interpretation. NR, BK, and KW jointly designed and supervised the study, and contributed to the interpretation of the data, and to the writing of the manuscript. All authors reviewed the final manuscript.

FUNDING

This study has received funding from the Federal Ministry of Food and Agriculture as part of the BlueCow Project (FKZ: 281B101516). The publication of this article was funded by the Open Access Fund of the FBN.

ACKNOWLEDGMENTS

We thank the entire team of Gut Dummerstorf GmbH, led by Lucas Pieper, for the excellent support in carrying out the trial. We thank Solvig Görs and Kirsten Kàrpàrti for the analysis of N-fractions in urine samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.699550/full#supplementary-material

REFERENCES

Abdoun, K., Stumpff, F., and Martens, H. (2006). Ammonia and urea transport across the rumen epithelium: a review. *Anim. Health Res. Rev.* 7, 43–59. doi: 10.1017/S1466252307001156

Aguilar, M., Hanigan, M. D., Tucker, H. A., Jones, B. L., Garbade, S. K., and Mcgilliard, M. L. (2012). Cow and herd variation in milk urea nitrogen concentrations in lactating dairy cattle. *J. Dairy Sci.* 95, 7261–7268. doi: 10.3168/jds.2012-5582

Balcells, J., Guada, J. A., Castrillo, C., and Gasa, J. (1991). Urinary excretion of allantoin and allantoin precursors by sheep after different rates of purine infusion into the duodenum. *J. Agric. Sci.* 116, 309–317.

Bouwman, A., Schopen, G., Bovenhuis, H., Visker, M., and Van Arendonk, J. (2010). Genome-wide scan to detect quantitative trait loci for milk urea

nitrogen in Dutch Holstein-Friesian cows. *J. Dairy Sci.* 93, 3310–3319. doi: 10.3168/jds.2009-2829

Brazert, M., Kranc, W., Nawrocki, M. J., Sujka-Kordowska, P., Konwerska, A., Jankowski, M., et al. (2020). New markers for regulation of transcription and macromolecule metabolic process in porcine oocytes during *in vitro* maturation. *Mol. Med. Rep.* 21, 1537–1551. doi: 10.3892/mmr.2020.10963

Bristow, A. W., Whitehead, D. C., and Cockburn, J. E. (1992). Nitrogenous constituents in the urine of cattle, sheep and goats. *J. Sci. Food Agric.* 59, 387–394. doi: 10.1002/jsfa.2740590316

Bryant, M. (1970). Normal flora—rumen bacteria. *Am. J. Clin. Nutr.* 23, 1440–1450. doi: 10.1093/ajcn/23.11.1440

Burgos, S. A., Embertson, N. M., Zhao, Y., Mitloehner, F. M., Depeters, E. J., and Fadel, J. G. (2010). Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: relation of milk urea nitrogen to ammonia emissions. *J. Dairy Sci.* 93, 2377–2386. doi: 10.3168/jds.2009-2415

Burgos, S. A., Fadel, J. G., and Depeters, E. J. (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. doi: 10.3168/jds.2007-0299

Cecchinato, A., Ribeca, C., Chessa, S., Cipolat-Gotet, C., Maretto, F., Casellas, J., et al. (2014). Candidate gene association analysis for milk yield, composition, urea nitrogen and somatic cell scores in Brown Swiss cows. *Animal* 8, 1062–1070. doi: 10.1017/S1751731114001098

Cox, C. (2013). *Experimentelle Untersuchungen zur Kinetik des Stickstoff-Umsatzes bei wachsenden Ziegenlämmern unter Verwendung des stabilen Isotops 15N.* [Dissertation]. Giessen: Justus Liebig University Giessen.

Farhadian, M., Rafat, S. A., Panahi, B., and Mayack, C. (2021). Weighted gene co-expression network analysis identifies modules and functionally enriched pathways in the lactation process. *Sci. Rep.* 11:2367. doi: 10.1038/s41598-021-81888-z

Fryer, M. W., Owen, V. J., Lamb, G. D., and Stephenson, D. G. (1995). Effects of creatine phosphate and P (i) on Ca2+ movements and tension development in rat skinned skeletal muscle fibres. *J. Physiol.* 482, 123–140. doi: 10.1113/jphysiol.1995.sp020504

Gao, X., Starmer, J., and Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* 32, 361–369. doi: 10.1002/gepi.20310

Gardiner, C., Clough, T. J., Cameron, K. C., Di, H. J., Edwards, G., and De Klein, C. (2016). Potential for forage diet manipulation in New Zealand pasture ecosystems to mitigate ruminant urine derived N2O emissions: a review. *N. Z. J. Agric. Res.* 59, 301–317. doi: 10.1080/00288233.2016.1190386

Gebreyesus, G., Buitenhuis, A. J., Poulsen, N. A., Visker, M., Zhang, Q., Van Valenberg, H. J. F., et al. (2019). Multi-population GWAS and enrichment analyses reveal novel genomic regions and promising candidate genes underlying bovine milk fatty acid composition. *BMC Genomics* 20:178. doi: 10.1186/s12864-019-5573-9

Gloerich, M., Ten Klooster, J. P., Vliem, M. J., Koorman, T., Zwartkruis, F. J., Clevers, H., et al. (2012). Rap2A links intestinal cell polarity to brush border formation. *Nat. Cell Biol.* 14, 793–801. doi: 10.1038/ncb2537

Gonzalez-Ronquillo, M., Balcells, J., Guada, J., and Vicente, F. (2003). Purine derivative excretion in dairy cows: endogenous excretion and the effect of exogenous nucleic acid supply. *J. Dairy Sci.* 86, 1282–1291. doi: 10.3168/jds.S0022-0302(03)73712-6

Graham, C., and Simmons, N. L. (2005). Functional organization of the bovine rumen epithelium. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 288, R173–R181. doi: 10.1152/ajpregu.00425.2004

Guarini, A. R., Lourenco, D.a.L, Brito, L. F., Sargolzaei, M., Baes, C. F., Miglior, F., et al. (2019). Genetics and genomics of reproductive disorders in Canadian Holstein cattle. *J. Dairy Sci.* 102, 1341–1353. doi: 10.3168/jds.2018-15038

Hale, S., Capuco, A., and Erdman, R. (2003). Milk yield and mammary growth effects due to increased milking frequency during early lactation. *J. Dairy Sci.* 86, 2061–2071. doi: 10.3168/jds.S0022-0302(03)73795-3

He, H., Chen, S., Liang, W., and Liu, X. (2017). Genome-wide proteomics analysis on longissimus muscles in Qinchuan beef cattle. *Anim. Genet.* 48, 131–140. doi: 10.1111/age.12508

Höfener, J. (2019). *Schätzung genetischer Parameter für den Milchharnstoffgehalt von Kühen der Rasse Deutsche Holstein-Schwarzbunt in konventionell und nach ökologischen Richtlinien wirtschaftenden Betrieben.* [Dissertation]. Giessen: Justus Liebig University Giessen.

Hopkins, P. C., Sáinz-Fuertes, R., and Lovestone, S. (2011). The impact of a novel apolipoprotein E and amyloid-β protein precursor-interacting protein on the production of amyloid-β. *J. Alzheimers Dis.* 26, 239–253. doi: 10.3233/JAD-2011-102115

Howard, A. D., Wang, R., Pong, S.-S., Mellin, T. N., Strack, A., Guan, X.-M., et al. (2000). Identification of receptors for neuromedin U and its role in feeding. *Nature* 406, 70–74. doi: 10.1038/35017610

Hu, Z. L., Park, C. A., and Reecy, J. M. (2019). Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* 47, D701–D710. doi: 10.1093/nar/gky1084

Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8:giy154. doi: 10.1093/gigascience/giy154

Jahnel, R. E., Wittenburg, D., Mayer, M., Täubert, H., and Reinsch, N. (2021). Estimation of genetic parameters of milk urea content. *Züchtungskunde* 93, 157–168.

Jardstedt, M., Hessle, A., Nørgaard, P., Richardt, W., and Nadeau, E. (2017). Feed intake and urinary excretion of nitrogen and purine derivatives in pregnant suckler cows fed alternative roughage-based diets. *Livest. Sci.* 202, 82–88. doi: 10.1016/j.livsci.2017.05.026

Jentsch, T. J., Stein, V., Weinreich, F., and Zdebik, A. A. (2002). Molecular structure and physiological function of chloride channels. *Physiol. Rev.* 82, 503–568. doi: 10.1152/physrev.00029.2001

Kehraus, S., Sudekum, K.-H., and Pfeffer, E. (2006). "Einflussfaktoren auf die ausscheidung N-haltiger verbindungen im harn von wiederkauern," in *Übersichten zur Tierernährung*, Vol. 34, ed. J. Kamphues (Frankfurt: DLG-Verlag), 125–164.

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

Kranc, W., Celichowski, P., Budna, J., Khozmi, R., Bryja, A., Ciesiółka, S., et al. (2017). Positive regulation of macromolecule metabolic process belongs to the main mechanisms crucial for porcine oocytes maturation. *Med. J. Cell Biol.* 5, 15–31. doi: 10.1515/acb-2017-0002

Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Gen.* 11, 241–247. doi: 10.1038/ng1195-241

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Liu, Y., Chen, X., Gong, Z., Zhang, H., Fei, F., Tang, X., et al. (2019). Fry is required for mammary gland development during pregnant periods and affects the morphology and growth of breast cancer cells. *Front. Oncol.* 9:1279. doi: 10.3389/fonc.2019.01279

Maas, J., France, J., and Mcbride, B. (1997). Model of milk protein synthesis. A mechanistic model of milk protein synthesis in the lactating bovine mammary gland. *J. Theor. Biol.* 187, 363–378. doi: 10.1006/jtbi.1997.0437

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

Müller, C. B. M., Görs, S., Derno, M., Tuchscherer, A., Wimmers, K., Zeyner, A., et al. (2021). Differences between Holstein dairy cows in renal clearance rate of urea affect milk urea concentration and the relationship between milk urea and urinary nitrogen excretion. *Sci. Total Environ.* 755:143198. doi: 10.1016/j.scitotenv.2020.143198

Munyaneza, N., Niyukuri, J., and Hachimi, Y. E. (2017). Milk urea nitrogen as an indicator of nitrogen metabolism efficiency in dairy cows: a review. *Theriogenol. Insight* 7, 145–159. doi: 10.5958/2277-3371.2017.00032.8

Neupane, M., Geary, T. W., Kiser, J. N., Burns, G. W., Hansen, P. J., Spencer, T. E., et al. (2017). Loci and pathways associated with uterine capacity for pregnancy and fertility in beef cattle. *PLoS One* 12:e0188997. doi: 10.1371/journal.pone.0188997

Pausch, H., Emmerling, R., Schwarzenbacher, H., and Fries, R. (2016). A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genet. Sel. Evol.* 48:14. doi: 10.1186/s12711-016-0190-4

Pegolo, S., Dadousis, C., Mach, N., Ramayo-Caldas, Y., Mele, M., Conte, G., et al. (2017). SNP co-association and network analyses identify E2F3, KDM5A and BACH2 as key regulators of the bovine milk fatty acid profile. *Sci. Rep.* 7:17317. doi: 10.1038/s41598-017-17434-7

Plante, I., and Laird, D. W. (2008). Decreased levels of connexin43 result in impaired development of the mammary gland in a mouse model of oculodentodigital dysplasia. *Dev. Biol.* 318, 312–322. doi: 10.1016/j.ydbio.2008.03.033

Pokharel, K., Peippo, J., Weldenegodguad, M., Honkatukia, M., Li, M. H., and Kantanen, J. (2020). Gene expression profiling of corpus luteum reveals important insights about early pregnancy in domestic sheep. *Genes* 11:415. doi: 10.3390/genes11040415

Rafiq, S., Huma, N., Pasha, I., Sameen, A., Mukhtar, O., and Khan, M. I. (2016). Chemical composition, nitrogen fractions and amino acids profile of milk from different animal pecies. *Asian Australas. J. Anim. Sci.* 29, 1022–1028. doi: 10.5713/ajas.15.0452

Rensing, S., Alkhoder, H., Kubitz, C., Schierenbeck, S., and Segelke, D. (2017). Cow reference population–benefit for genomic evaluation system and farmers. *Interbull Bull.* 51, 63–66.

Ribeiro, E. S., Monteiro, A. P. A., Bisinotto, R. S., Lima, F. S., Greco, L. F., Ealy, A. D., et al. (2016). Conceptus development and transcriptome at preimplantation stages in lactating dairy cows of distinct genetic groups and estrous cyclic statuses. *J. Dairy Sci.* 99, 4761–4777. doi: 10.3168/jds.2015-10315

Roginski, H., Fuquay, J. W., and Fox, P. F. (2003). *Encyclopedia of Dairy Sciences*. Cambridge, MA: Academic press.

Saatchi, M., Schnabel, R. D., Taylor, J. F., and Garrick, D. J. (2014). Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics* 15:442. doi: 10.1186/1471-2164-15-442

Sampson, C. M., Kasper, J. M., Felsing, D. E., Raval, S. R., Ye, N., Wang, P., et al. (2018). Small-molecule neuromedin U receptor 2 agonists suppress food intake and decrease visceral fat in animal models. *Pharmacol. Res. Perspect.* 6:e00425. doi: 10.1002/prp2.425

Sargolzaei, M., Chesnais, J. P., and Schenkel, F. S. (2010). "Accuracy of a family-based genotype imputation algorithm," in *Proceedings of the 32nd Conference for the International Society for Animal Genetics*, (Edinburgh: ISAG).

Sawa, A., Bogucki, M., and Krêñel-Czopek, S. (2011). Effect of some factors on relationships between milk urea levels and cow fertility. *Arch. Anim. Breed.* 54, 468–476. doi: 10.5194/aab-54-468-2011

Schiaffino, M. C., Bellini, C., Costabello, L., Caruso, U., Jakobs, C., Salomons, G. S., et al. (2005). X-linked creatine transporter deficiency: clinical description of a patient with a novel SLC6A8 gene mutation. *Neurogenetics* 6, 165–168. doi: 10.1007/s10048-005-0002-4

Schwiebert, E. M., Lopes, A. G., and Guggino, W. B. (1994). Chloride channels along the nephron. *Curr. Top. Membr.* 42, 265–315. doi: 10.1016/S0070-2161(08)60825-4

Sebek, L., Van Riel, J., and De Jong, G. (2007). *The Breeding Value for Milk Urea as Predictor for the Efficiency of Protein Utilization in Dairy Cows.* Report 81 from Animal Sciences Group of Wageningen University. Wageningen: Wageningen University.

Spek, J. W., Dijkstra, J., Van Duinkerken, G., and Bannink, A. (2012). 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. doi: 10.1017/s0021859612000561

Stamer, E., Brade, W., and Thaller, G. (2011). Modellentwicklung und Schätzung genetischer parameter für den Harnstoffgehalt in der Milch erst-und zweitlaktierender Holstein-Kühe. *Züchtungskunde* 83, 104–117.

Stefanon, B., Colitti, M., Gabai, G., Knight, C. H., and Wilde, C. J. (2002). Mammary apoptosis and lactation persistency in dairy animals. *J. Dairy Res.* 69, 37–52. doi: 10.1017/S0022029901005246

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

Sun, H. Z., Zhou, M., Wang, O., Chen, Y., Liu, J. X., and Guan, L. L. (2020). Multi-omics reveals functional genomic and metabolic mechanisms of milk production and quality in dairy cows. *Bioinformatics* 36, 2530–2537. doi: 10.1093/bioinformatics/btz951

Susmel, P., Spanghero, M., Stefanon, B., and Mills, C. (1995). Nitrogen balance and partitioning of some nitrogen catabolites in milk and urine of lactating cows. *Livest. Prod. Sci.* 44, 207–219. doi: 10.1016/0301-6226(95)00074-7

Tas, B. M., and Susenbeth, A. (2007). Urinary purine derivates excretion as an indicator of in vivo microbial N flow in cattle: a review. *Livest. Sci.* 111, 181–192. doi: 10.1016/j.livsci.2007.05.010

Uwizeye, A., De Boer, I. J. M., Opio, C. I., Schulte, R. P. O., Falcucci, A., Tempio, G., et al. (2020). Nitrogen emissions along global livestock supply chains. *Nat. Food* 1, 437–446. doi: 10.1038/s43016-020-0113-y

Vekaria, R. M., Shirley, D. G., Sevigny, J., and Unwin, R. J. (2006). Immunolocalization of ectonucleotidases along the rat nephron. *Am. J. Physiol. Renal Physiol.* 290, F550–F560. doi: 10.1152/ajprenal.00151.2005

Verrey, F., Closs, E. I., Wagner, C. A., Palacin, M., Endou, H., and Kanai, Y. (2004). CATs and HATs: the SLC7 family of amino acid transporters. *Pflugers Arch. Eur. J. Physiol.* 447, 532–542. doi: 10.1007/s00424-003-1086-z

Vijayakumar, P., Bakyaraj, S., Singaravadivelan, A., Vasanthakumar, T., and Suresh, R. (2019). Meta-analysis of mammary RNA seq datasets reveals the molecular understanding of bovine lactation biology. *Genome* 62, 489–501. doi: 10.1139/gen-2018-0144

Wang, L., Chang, J. H., Paik, S. Y., Tang, Y., Eisner, W., and Spurney, R. F. (2011). Calcineurin (CN) activation promotes apoptosis of glomerular podocytes both *in vitro* and *in vivo*. *Mol. Endocrinol.* 25, 1376–1386. doi: 10.1210/me.2011-0029

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

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

11.2 Rumen microbiota and host gene expression associate to predisposed milk urea (MU) concentration in Holsteins (2nd study)

Rumen microbiota and host gene expression associate to predisposed milk urea (MU) concentration in Holsteins.

H. Honerlagen[1*], H. Reyer[1], D. Segelke[2], M. Oster[1], S. Ponsuksili[1], N. Trakooljul[1], B. Kuhla[1], K. Wimmers[1,3]

[1] Research Institute for Farm Animal Biology (FBN), Wilhelm-Stahl-Allee 2, 18196, Dummerstorf, Germany; [2] IT-Solutions for Animal Production (vit), Heinrich-Schröder-Weg 1, 27283 Verden, Germany; [3] University of Rostock, Justus-von-Liebig-Weg 6, 18059, Rostock, Germany; *honerlagen@fbn-dummerstorf.de

Abstract

The dairy sector has to reduce nitrogen (N) surpluses in the near future. In this context, the cow individual milk urea (MU) concentration has been suggested to reflect N-excretion in breeding strategies. The variability in MU levels is partly affected by ruminal processes attributed to microbes and host gene expression. To investigate the relevance of MU for breeding, we aimed to uncover ruminal mechanisms in cows with predisposed higher compared to lower MU (HMU vs. LMU; n=20). Whereas the core microbiota at phylum level remained constant, the abundances of three microbial families differed between HMU and LMU cows. Transcriptome profiling of rumen tissue revealed 28 genes, including *MMP3*, *TNP2* and various genes of immune response that distinguished HMU and LMU cows. Our findings contribute to a deeper knowledge of MU predispositions in Holsteins and attribute to the optimisation of N-usage in conjunction with the reduction of N-emissions from dairy farms.

Introduction

The primary source of N-emissions from dairy cows is displayed by unutilised N, which is ingested by the cow with feed crude protein and excreted via urine, milk and faeces (Abdoun *et al*., 2006). In this context, the ruminal conversion rate of N into valuable microbial protein, which is the result of rumen microbial growth, considerably contributes to the amount of N-utilisation and N-emissions (Tan *et al*., 2021). Due to the high phenotypic correlation between milk urea (MU), urinary urea and blood urea concentration (Guliński *et al*., 2016) the cow individual MU value has been suggested as substitute trait for breeding selection (Beatson *et al*., 2019; Honerlagen *et al*., 2021). In regard of successful breeding selection on N-use efficiency and reduced N-excretion from dairy cows, unravelling the molecular biological background of predisposed higher or lower MU concentrations (HMU – LMU) in dairy cows is required. Considering the large impact of ruminal processes on the N-metabolism in lactating ruminants, our study elucidated exclusively differences in the rumen microbial community and corresponding host-specific transcriptional processes between HMU and LMU cows.

Materials & Methods

The predisposition for higher or lower MU concentration (HMU – LMU) was determined by deregressed proofs of a breeding value for MU, estimated by vit Verden (Vereinigte Informationssysteme Tierhaltung, Verden, Germany). The subsequent categorisation was further validated by the five recent milk records from each cow (HMU=278.1 ± 15.8 mg/l, LMU 181.54 ± 15.08 mg/l; mean ± SD). Based on this classification, 20 German Holsteins in late lactation were stalled in pairs, comprising one HMU and one LMU cow, treated as previously described (Müller *et al*., 2021) and slaughtered after two weeks. Samples of rumen villi and rumen fluids (50 ml) were collected from each cow immediately after slaughter, snap frozen (-80°C) and stored.

Laboratory analyses and sequencing pipelines. Microbial DNA was extracted from rumen fluids, quality checked and used to amplify the V4 hypervariable region of the 16S rRNA gene (Kozich *et al*., 2013), which was sequenced on HiSeq2500 (Illumina, San Diego, CA). The reads were subsequently filtered, trimmed, and removed from chimeras within the mothur software (version 1.44.1; Schloss *et al*., 2009). Sequences were aligned to the Silva reference database (release 138; https://www.arb-silva.de/), aggregated into operational taxonomic units (OTU, sequence identity ≥97%) and annotated according to the Silva database. The count data was rarefied to the sample with the lowest read depth (745,235 reads). Statistical analyses were conducted at phylum (n=19) and family (n=116) levels. Total RNA was isolated from homogenized rumen villus epithelium, purified, qualitatively and quantitatively checked and utilized for mRNA library preparation. Libraries were paired-end sequenced on HiSeq2500. After quality control and filtering, sequences were aligned to the current reference genome (Bos_taurus.ARS-UCD1.2). The number of reads uniquely mapped to each gene were extracted using HTSeq (version 0.12.4).

Statistical Analyses of Rumen Microbiome and Transcriptome. Initially the core microbiome of HMU and LMU cows was investigated based on relative abundances at phylum level. Subsequently, DESeq2 was used to identify differentially abundant phyla and families between HMU and LMU cows (Love *et al*., 2014). Similarly, the transcriptome dataset was analysed for differentially expressed genes (DEG) with DESeq2. The trial of each pair was included in all statistical models as fixed effect accounting for different crude protein levels in the feed rations. DEG were further assigned to their general biological functions using GeneCards (http://www.genecards.org).

Results

Bacteroidota und *Firmicutes* displayed the most dominant phyla in both, HMU and LMU cows by capturing 78.7% of relative abundances in all rumen fluids (43.4% and 35.3%, respectively) (Figure 1). The analysis at family level revealed *Monoglobaceae* ($p<0.01$) and *Ruminococcaceae* ($p=0.01$) being significantly more abundant in LMU compared to HMU cows. Microbes assigned to *Acetobacteraceae* sowed a significantly high relative occurrence in HMU compared to LMU cows ($p=0.04$).

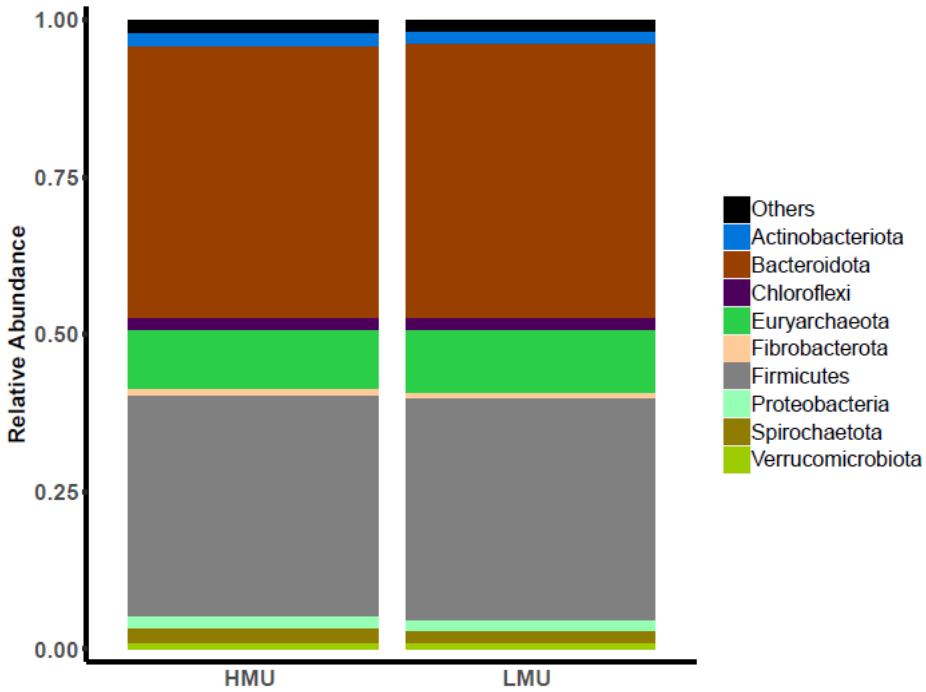

Figure 1. Microbial composition of rumen fluids from HMU and LMU cows at phylum level indicating a similar pattern with dominance of microbes assigned to *Bacteroidota* and *Firmicutes*.

In the filtered transcriptome dataset 28 DEG were identified between HMU and LMU animals. The highest fold changes were revealed for Matrix Metallopeptidase 3 (*MMP3*) and Transition Protein 2 (*TNP2*), with higher transcript abundance in HMU and LMU, respectively (Table 1). Five of the top ten DEG were identified as contributors of immune response and displayed higher expression in LMU cows.

Table 1. Most prominent DEG between HMU and LMU cows.

Gene Symbol	Log 2 fold change[1]	Adjusted *p*-value[2]	General function[3]
TNP2	-1.40	0.0159	Fertility
MMP3	1.22	0.0356	Tissue remodeling
OAS1Y	-1.18	0.0316	Immune system
MX1	-1.09	0.0266	Immune system
IRF4	-0.99	0.0266	Immune system
PLA2G2D4	-0.95	0.0208	Immune system
SLAMF1	-0.94	0.0145	Immune system
GALNT8	0.91	0.0045	Protein metabolism
GPRC5A	0.91	0.0273	Cell development
CHI3L2	-0.84	0.0216	Cartilage biogenesis

[1] derived from DESeq2 analysis and refer to HMU
[2] Benjamini-Hochberg adjusted *p*-values
[3] General functions were obtained from GeneCards database

Discussion

Our study focused on the ruminal characterisation of Holsteins with predisposed higher or lower MU concentration (HMU – LMU). Both, HMU and LMU cows occupied the typical core microbiome of the bovine rumen (Wallace *et al*., 2019) comprising mainly *Bacteroidota* and *Firmicutes*, which captured 78.7% of total rumen abundances (43.4% and 35.3% respectively). Apparently, MU predisposition was not substantially associated with distinguished rumen flora. However, three significantly differentially abundant families between HMU and LMU cows were identified. Interestingly, the *Ruminococcaceae* family, which displayed higher abundances in LMU cows has already been attributed a fundamental role in amino acidic and protein metabolism in the bovine rumen (Pacífico *et al*., 2021). In addition, various *Ruminococcaceae* genera displayed higher abundances in goats with high N-efficient phenotypes (Wang *et al*., 2019). Further analysis of the microbiome at genus level as well as the effect estimation of specific genera on N-utilisation and N-excretion might unravel substantial information of ruminal MU characterisation. The results of DEG analysis elucidated the enhanced expression of immune response genes in LMU cows. Interestingly, Zhang *et al*. reported the activation of immune response pathways in the rumen epithelium of calves with higher N-efficiency (Zhang *et al*., 2021). Moreover, recent studies identified immune system activities of the host epithelia as key drivers of long- term host – microbial maintenance (Peterson *et al*., 2007). In this context, the interplay between host gene expression, the microbial community and the rumen environment as well as the different partitioning of nitrogen to body compartments will be further investigated between HMU and LMU cows with regard to divergent N utilisation and N excretion mechanisms.

References

Abdoun K., Stumpff F., and Martens H. (2006) Anim Health Res Rev 7(1-2)**:**43-59. https://doi.org/10.1017/S1466252307001156

Beatson P. R., Meier S., Cullen N. G., Eding, H. (2019) Animal, 13(10): 2164-2171. https://doi.org/10.1017/S1751731119000235

Guliński P., Salamończyk E., and Młynek K. (2016) Anim Sci Pap Rep 34(1):5-23.

Honerlagen H., Reyer H., Oster M., Ponsuksili S., Trakooljul N., *et al*. (2021) Front Genet 12:1241. https://doi.org/10.3389/fgene.2021.699550

Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K., and Schloss P.D. (2013) Appl Environ Microbiol 79(17):5112-5120. https://doi.org/10.1128/AEM.01043-13

Love M.I., Huber W., and Anders S. (2014) Genome Biol 15(12):1-21. https://doi.org/10.1186/s13059-014-0550-8

Müller C.B.M., Görs S., Derno M., Tuchscherer A., Wimmers K., *et al*. (2021) Sci Total Environ 755(2):143-198. https://doi.org/10.1016/j.scitotenv.2020.143198

Pacífico C., Petri R.M., Ricci S., Mickdam E., Wetzels S.U., *et al*. (2021) Microorganisms 9(2):342. https://doi.org/10.3390/microorganisms9020342

Peterson D.A., Mcnulty N.P., Guruge J.L., and Gordon J.I. (2007) Cell Host Microbe 2(5):328-339. https://doi.org/10.1016/j.chom.2007.09.013

Tan P., Liu H., Zhao J., Gu X., Wei X., *et al.* (2021) Sci Total Environ 800:149-596. https://doi.org/10.1016/j.scitotenv.2021.149596

Wallace R.J., Sasson G., Garnsworthy P.C., Tapio I., Gregson E., *et al*. (2019) Sci Adv 5(7):83-91. https://doi.org/10.1126/sciadv.aav8391

Wang L., Liu K., Wang Z., Bai X., Peng Q., *et al*. (2019) Front Microbiol 10:239. https://doi.org/10.3389/fmicb.2019.00239

Zhang N., Teng Z., Li P., Fu T., Lian H., *et al*. (2021) PloS one 16(9):257-417. https://doi.org/10.1371/journal.pone.0257417

11.3 Ruminal background of predisposed milk urea (MU) concentration in Holsteins (3rd study)

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 13 September 2022
DOI 10.3389/fmicb.2022.939711

OPEN ACCESS

EDITED BY
Alejandro Belanche,
University of Zaragoza, Spain

REVIEWED BY
Francesco Rubino,
Queen's University Belfast,
United Kingdom
Juan Manuel Palma-Hidalgo,
Scotland's Rural College,
United Kingdom

*CORRESPONDENCE
Klaus Wimmers
wimmers@fbn-dummerstorf.de

SPECIALTY SECTION
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

RECEIVED 09 May 2022
ACCEPTED 26 July 2022
PUBLISHED 13 September 2022

CITATION
Honerlagen H, Reyer H, Segelke D, Müller CBM, Prahl MC, Ponsuksili S, Trakooljul N, Reinsch N, Kuhla B and Wimmers K (2022) Ruminal background of predisposed milk urea (MU) concentration in Holsteins. *Front. Microbiol.* 13:939711. doi: 10.3389/fmicb.2022.939711

Ruminal background of predisposed milk urea (MU) concentration in Holsteins

Hanne Honerlagen[1], Henry Reyer[1], Dierck Segelke[2], Carolin Beatrix Maria Müller[3], Marie Christin Prahl[3], Siriluck Ponsuksili[1], Nares Trakooljul[1], Norbert Reinsch[4], Björn Kuhla[3] and Klaus Wimmers[1,5]*

[1]Research Institute for Farm Animal Biology (FBN), Institute of Genome Biology, Dummerstorf, Germany, [2]IT-Solutions for Animal Production, Vereinigte Informationssysteme Tierhaltung w.V. (vit), Verden, Germany, [3]Research Institute for Farm Animal Biology (FBN), Institute of Nutritional Physiology "Oskar Kellner", Dummerstorf, Germany, [4]Research Institute for Farm Animal Biology (FBN), Institute of Genetics and Biometry, Dummerstorf, Germany, [5]Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

Efforts to reduce nitrogen (N) emissions are currently based on the optimization of dietary- N supply at average herd N requirements. The implementation of the considerable individual differences and predispositions in N- use efficiency and N- excretion in breeding programs is hampered by the difficulty of data collection. Cow individual milk urea (MU) concentration has been proposed as an easy-to-measure surrogate trait, but recent studies questioned its predictive power. Therefore, a deeper understanding of the biological mechanisms underlying predisposed higher (HMUg) or lower (LMUg) MU concentration in dairy cows is needed. Considering the complex N- metabolism in ruminants, the distinction between HMUg and LMUg could be based on differences in (i) the rumen microbial community, (ii) the host-specific transcription processes in the rumen villi, and (iii) the host–microbe interaction in the rumen. Therefore, rumen fluid and rumen epithelial samples from 10 HMUg and 10 LMUg cows were analyzed by 16S sequencing and HiSeq sequencing. In addition, the effect of dietary-N reduction on ruminal shifts was investigated in a second step. In total, 10 differentially abundant genera (DAG) were identified between HMUg and LMUg cows, elucidating greater abundances of ureolytic *Succinivibrionaceae_UCG-002* and *Ruminococcaceae_unclassified* in LMUg animals and enhanced occurrences of *Butyvibrio* in HMUg cows. Differential expression analysis revealed genes of the bovine Major Histocompatibility Complex (*BOLA* genes) as well as *MX1*, *ISG15*, and *PRSS2* displaying candidates of MU predisposition that further attributed to enhanced immune system activities in LMUg cows. A number of significant correlations between microbial genera and host transcript abundances were uncovered, including strikingly positive correlations of *BOLA-DRA* transcripts with *Roseburia* and *Lachnospiraceae* family abundances that might constitute particularly prominent microbial–host interplays of MU predisposition. The reduction of feed-N was followed by 18 DAG in HMUg and 19 DAG in LMUg, depicting pronounced interest on *Shuttleworthia*, which displayed controversial adaption in HMUg and LMUg cows. Lowering feed-N further elicited massive downregulation of immune

response and energy metabolism pathways in LMUg. Considering breeding selection strategies, this study attributed information content to MU about predisposed ruminal N-utilization in Holstein–Friesians.

KEYWORDS

rumen microbiome, host gene expression, microbe–host relationship, dairy cow, milk urea

Introduction

Nitrogen (N) excretions of dairy cows contribute substantially to N-deposition in the environment (Castillo et al., 2000). Considering the relevance for climate change (nitrous oxide (N_2O) atmospheric emissions) and groundwater quality (nitrate (NO_3^-) contamination), the dairy sector is demanded to reduce its N-emissions in the near future (Uwizeye et al., 2020). The primary source of N-emissions from dairy cows is represented by unutilized N, which goes into the cow *via* feed crude protein (CP) and is excreted *via* urine, milk and feces (Bryant, 1970; Abdoun et al., 2006). In this context, the ruminal conversion rate of feed N into valuable microbial protein, which is the result of rumen microbial growth, mainly determines the number of N-emissions (Tan et al., 2021).

In detail, the cow acquires N *via* the dietary CP. A small part of CP passes through the rumen as undegradable protein (UDP) and is absorbed in the small intestine or excreted *via* feces. The majority of dietary N from CP is provided to the rumen microbes for microbial growth. However, the rate of microbial N-utilization depends on various factors such as ruminal pH, dietary energy content, the absorption capacity of the rumen epithelium and the microbial community. Thus, a part of the ruminal N-influx is incorporated into microbial protein, whereas the remaining part is microbially converted to non-protein-N (NPN) in forms of ammonia (NH_3), ammonium-ions (NH_4^+), peptides and amino acids (Bryant, 1970; Hartinger et al., 2018; Cholewińska et al., 2020). The cow absorbs amino acids and peptides deriving from the microbial protein in the small intestine whereas the NPN-compounds are absorbed by the rumen epithelium. The amino acids and peptides can be utilized by the cow as valuable N-sources, while absorbed NH_3 and NH_4^+ is transported with the blood stream to the liver. There, the ammonia is detoxified, yielding urea in an energy-intensive process. Urea is secreted back into the blood cycle from which it is eliminated *via* urine (urinary urea, UU) and milk (milk urea, MU), or secreted into the rumen either across the rumen epithelium or *via* saliva (Burgos et al., 2007; Spek et al., 2012). The recycled ruminal urea is converted by epithelium-bound, ureolytic rumen bacteria to CO_2 and NH_3. The latter, in turn, can be utilized by NH_3-assimilating microbes for the metabolism of microbial protein (Jin et al., 2018).

In symbiosis with their microbes, dairy cows may utilize on average 25% of the dietary N (Calsamiglia et al., 2010). However, substantial variation in the N-utilization efficiency, which is defined as the ratio of gram N in product to gram N intake between individual cows has been observed, ranging from 15 to 40% (Calsamiglia et al., 2010; Spek et al., 2012; Gardiner et al., 2016; Chen et al., 2021). In recent years various studies have focused on the reduction of dietary CP levels to minimize N-emissions at herd level, but neither the cow-specific potential has been exploited intensely nor has the predisposition for individual N-use efficiency been considered for breeding progress yet (Cantalapiedra-Hijar et al., 2020; Bergen, 2021; Jahnel et al., 2021; Te Pas et al., 2021). One reason is that data collection for the cow-individual N-use efficiency is not feasible in huge cow cohorts. Generally, neither the cow individual N-uptake nor the content of N in urine and feces is measurable in practice operating dairy farms. Thus, in the last years the cow-specific MU concentration has been considered as a substitute trait (Burgos et al., 2007, 2010; König et al., 2008; Honerlagen et al., 2021; Jahnel et al., 2021). MU values are available to breeding companies at a large scale due to the results of monthly milk records. Furthermore, MU is moderately heritable and only weak or even no genetic correlations to further milk performance parameters (i.e., milk protein yield, milk fat yield) were evidenced (Wood et al., 2003; Miglior et al., 2007), which would, in principle, support MU as a trait for breeding selection. Due to the phenotypic correlation between MU, UU and blood urea concentration (Spek et al., 2012; Guliński et al., 2016) it has been hypothesized that breeding selection on lower MU concentration could lead to lower N-urine concentration and thus to lower N-emissions from dairy farming (Beatson et al., 2019). In addition, lower MU concentrations were also suggested as biomarkers for increased N-utilization efficiency (Nousiainen et al., 2004; Lavery and Ferris, 2021), being predominantly explained by enhanced ruminal N-utilization (Guliński et al., 2016). However, up to now the validity of MU as a substitute trait for cow-individual N-use efficiency and N-excretion is relying on the phenotypic correlation between MU and UU, but its utility has been questioned by the results of two recent studies (Correa-Luna et al., 2021; Müller et al., 2021). Thus, it became evident, that a deeper knowledge of the biological mechanisms underlying

predisposed MU concentrations is indispensable before MU can be seriously proposed for breeding selection strategies.

Considering the large impact of ruminal processes on the N-metabolism in dairy cows, this study aims to elucidate ruminal mechanisms that distinguish between cows with predisposed higher (HMUg) and lower (LMUg) MU concentration to contribute to the validation of MU as substitutive trait for N-utilization efficiency and N-excretion. In detail, this study focuses on differences in (i) the rumen microbial community, (ii) the host-specific transcriptional processes in the rumen villi and (iii) the host–microbe interaction in the rumen between HMUg and LMUg cows.

Moreover, assumed that LMUg cows may possess lower N-input requirements as a consequence of more efficient N-utilization, which would result in an N-oversupply if fed a ration according to the recommendations, this study further investigated the effect of reduced dietary N-supply on the microbial community and on host specific rumen tissue processes in HMUg and LMUg cows.

Materials and methods

Animal husbandry and sampling were carried out according to the guidelines of the German Animal Protection Law. All protocols were approved by the Institute's Animal Welfare Commission. The experimental protocol is in strict compliance with the German Animal Welfare Legislation, and has been approved by the Ethics Committee of the Federal State of Mecklenburg, Western Pomerania, Germany (Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei; LALLF M-V7221.3-2-019/19) and is in accordance with the ARRIVE guidelines.

Cow population and sample collection

The predisposition of cows for a higher or lower MU concentration (HMUg and LMUg) was determined by deregressed proofs of an estimated breeding value for MU (EBVMU). The EBVMUs were estimated by vit Verden (Vereinigte Informationssysteme Tierhaltung, Verden, Germany) applying the established model for somatic cell score in milk [Vereinigte Informationssysteme Tierhaltung (VIT)., (2020, access 20.01.2020)]. The estimation was based on a milk record dataset of around 8 million lactation curves of milk urea values from Holstein cows in their first to third parities. The deregression of EBVMUs is a pre step for genomic breeding value estimation. The HMUg was determined as predisposition by positive values, LMUg by negative values. Based on this categorization, 10 HMUg and 10 LMUg cows were housed in pairs, comprising one HMUg and one LMUg cow, respectiveley. The phenotypes of MU concentration were obtained from the recent five milk records depicting HMUg = 278.1 ± 15.8 (mean ± SD) mg/L and LMUg = 181.54 ± 15.08 (mean ± SD) mg/L with comparable milk yields between the groups.

These cow groups were fed an isocaloric diet [10.1 ± 0.2 MJ metabolisable energy/kg dry matter (DM)], with either recommended (normal) CP content (NP with 157 ± 2.68 (mean ± SD) g CP/kg DM) or low CP content [LP with 139 ± 8.37 (mean ± SD) g CP/kg DM], resulting in a 4 × 5 balanced design (NPxHMUg, NPxLMUg, LPxHMUg, LPxLMUg; each n = 5). The experiment was conducted in 10 trials, with each trial comprising one cow pair (HMUg and LMUg), which was fed either the NP or the LP diet. The LP and NP rations were offered twice-daily for *ad libitum* intake as a total mixed ration. A detailed description of the diet, the feeding scheme as well as phenotypic analyses of N-metabolites (i.e., ammonia concentration in the rumen fluids, plasma urea, and glutamine concentration) is documented in Müller et al. (2021). After 2 weeks, the cows were slaughtered 4 h after morning milking and feeding by exsanguination after captive bolt stunning. Samples of rumen villi from the *saccus dorsalis* containing all rumen villus cell layers were collected, i.e., the basal membrane cells, which connect villi to blood stream and muscle tissue. The samples were rinsed, snap frozen in liquid nitrogen and stored at −80°C until RNA isolation. In addition, the rumen fluids (50 ml) were collected as a homogeneous mixture of the entire rumen content and stored at −20° until isolation of microbial DNA.

Rumen fluids: Microbial DNA extraction, 16S rRNA amplicon sequencing, and data preparation

The rumen fluid samples (50 ml) were mixed and microbial DNA was extracted with the PowerLyzer PowerSoil DNA isolation kit (QIAGEN, Hilden, Germany) following manufacture's recommendations, but utilizing 800 µl rumen fluid sample (instead of 250 µl) and two additional heat incubation steps (step 1: 70°C, 10 min; step 2: 95°C, 10 min) before the bead beating procedure. The quantity of the extracted DNA was determined on NanoDrop ND-2000 (Thermo Fisher Scientific, Dreieich, Germany) and afterwards diluted to 10 ng/µl per sample. The diluted DNA extracts were utilized to amplify the V4 hyper-variable region of the 16S rRNA gene. The specific primer enclosed the V4-specific sequence, the sequencing flow cell adapter and a specific barcode for each primer as previously described (Kozich et al., 2013). The amplification was conducted in duplicate by polymerase chain reaction (PCR) with initial denaturation at 95°C for 2 min, followed by 35 cycles at

95°C for 30 s, 50°C for 60 s, and 72°C for 90 s, and a final extension for 10 min at 72°C using the 5 Prime HotMasterMix (5 Prime, Hamburg, Germany). The PCR products (16S amplicons) were checked on agarose gel and subsequently purified, normalized and pooled for sequencing using the SequalPrep normalization plate kit (Thermo Fisher Scientific, Dreieich, Germany) The sequencing of the amplicons was conducted on HiSeq2500 (Illumina, San Diego, CA) with 250 bp paired-end reads.

The sequence data were then trimmed for the adapter sequences and filtered with mothur software for sequences that exceeded 275 bp in length, which contained ambiguous base calls or excessively long homopolymers (>8 bases), as well as putative chimeras were removed [version 1.44.1 (Schloss et al., 2009)].The sequences were further globally aligned to the Silva reference database (release 138; https://www.arb-silva.de/, access 10.6.20) and subsequently aggregated into operational taxonomic units (OTU) considering a sequence identity of ≥97%. Taxonomic annotations for OTUs were obtained from the Silva database (release 138). Considering stratification by dispersion, the count data was rarefied to the sample with the lowest read depth (745,235 reads) as suggested by Mcmurdie and Holmes (2014).

Rumen fluids: Statistical analysis of 16S rRNA sequencing data

Initially, very low abundant OTUs with <10 counts in all 20 cows were pre-discarded from further analyses. The microbial community in rumen fluids was analyzed with inverse Simpson diversity index (alpha diversity) at OTU level utilizing the R packages "vegan" (Oksanen et al., 2013) and "agricolae" (De Mendiburu Delgado, 2009). Means of Simpson indices of the predisposition groups (HMUg and LMUg) and the diet groups (NP and LP in HMUg and LMUg cows, respectively) were tested for significant differences by applying Student's *t*-test. The subsequent analyses were conducted at genus level, excluding genera with <10 counts in 5 or more cows to avoid potential noise by barely observed reads. The remaining dataset was tested for significantly differentially abundant genera (DAG) between the groups (HMUg–LMUg; NPxHMUg–LPxHMUg; NPxLMUg–LPxLMUg) utilizing Wald test in the DESeq2 R package (Love et al., 2014). For the contrast of predisposition (HMUg–LMUg) the trial was implemented as further fixed effect in the statistical model, accounting for different CP levels in the feed ration. For the differences between the diets, the contrasts within each predisposition group were analyzed and expressed as fold changes obtained from DESeq2 delogarithmized log2fold changes. Genera with *p*-values below 0.05 were considered as DAG.

Rumen tissue: RNA extraction, library preparation, sequencing, and data preparation

The total RNA from homogenized rumen villi was isolated using TRI reagent (Sigma–Aldrich, Taufkirchen, Germany), followed by DNaseI treatment (Roche, Mannheim, Germany) and a subsequent column-based purification step utilizing the NucleoSpin RNA Kit (Macherey-Nagel, Düren, Germany). The RNA quantity was measured on NanoDrop ND-2000 (Thermo Fisher Scientific) and further assessed for integrity by application on agarose gel and *via* 2100 Bioanalyzer system (Agilent, Santa Clara, CA). The RNA library preparation was performed using the TruSeq Stranded mRNA kit according to the manufacture's recommendation (Illumina). The resulting libraries were quality checked on the 2100 Bioanalyzer system and quantified using the Invitrogen Qubit dsDNA HS kit (Thermo Fisher Scientific). Libraries were paired-end sequenced for 2 × 71 cycles on HiSeq 2500 (Illumina). The raw sequencing reads (fastq) were quality-checked utilizing FastQC, version 0.11.8 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The data preprocessing was conducted with Trim Galore v.0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The low-quality reads (mean Q-score < 20) and short reads (<20 bp) were filtered out. Adapter-like sequences at the 3'-end of sequence reads were trimmed. The remaining high quality paired-end reads comprised a read depth of 20,324,785 ± 3,234,163 (mean ± SD) per sample and were further aligned to the current reference genome, Bos_taurus.ARS-UCD1.2 (access 01. 08.2020) with an average alignment rate of 97.86% ± 0.01% (SD) using Hisat2, version 2.2.0 (Kim et al., 2015, 2019; Pertea et al., 2016). The number of reads uniquely mapped to each gene were extracted from the HISAT2 mapping results using HTSeq, version 0.12.4 (Anders et al., 2015).

Rumen tissue: Statistical analysis of RNAseq data

The RNA dataset was analyzed for significantly differentially expressed genes (DEG) between groups—corresponding to the 16S microbiota sequence data (HMUg–LMUg; NPxHMUg–LPxHMUg; and NPxLMUg–LPxLMUg)—by utilizing Wald test in DESeq2 package (Love et al., 2014). Genes with <10 counts in 10 or more cows were excluded from analysis, leading to a dataset of 12,627 genes. The trial was considered as further fixed effect for HMUg–LMUg contrast. Genes with a Benjamini–Hochberg adjusted *p*-values <0.05 were considered as DEG. The DEG identifiers were extracted to investigate the biological function of the respective gene using the databases GeneCards (http://www.genecards.org), Ensembl (http://www.

ensembl.org), and the published literature. Furthermore, genes revealing *p*-values below 0.01 (corresponding adjusted *p*-values <0.6) were imported into Ingenuity Pathway Analysis (IPA; Ingenuity® Systems, http://www.ingenuity.com) for enrichment analysis between the groups (HMUg–LMUg; NPxHMUg–LPxHMUg; NPxLMUg–LPxLMUg). Pathways with an adjusted *p*-value <0.05 were defined as significantly enriched between the groups. Significant pathway inhibition or activation was defined by *Z*-scores (significant inhibition: *Z*-score <-2; significant activation: *Z*-score >2) in IPA. The *Z*-scores between −2 and 0 or 0 and 2 were considered indicators of inhibition or activation, respectively. The pathways were further assigned to general functions after evaluation in Targeted Explorer (IPA). The pathways of cancer and disease were excluded.

Integration of microbial and transcriptome data: Statistical analysis

Datasets of rumen microbial genera ($n = 229$ microbes) and gene expression ($n = 12{,}627$ genes) were initially transformed by variance-stabilizing transformation (DESeq2 R package). Subsequently, sparse Partial Least Squares Discriminant Analysis (sPLS-DA) was applied on both datasets. The sPLS-DA is incorporated in R package "mixOmics" (version 6.6.2, available at: http://mixomics.org) (Rohart et al., 2017) and achieves feature selection and dimension reduction simultaneously in dependence on a specific trait (Chun and Keleş, 2010). In this case, sPLS-DA was utilized to select a key subset of microbial genera and a key subset of genes that discriminate the microbial community and the rumen tissue transcriptome between HMUg and LMUg cows on two components each. The discrimination by the generated subsets was visualized in two plots.

The data subsets were further analyzed for microbial abundance–gene transcript correlations by utilizing Pearson statistics. Microbe–gene pairs exhibiting *p*-values below 0.01 were considered as significantly correlated. The results were visualized in a correlation map using R package "pheatmap" (version 1.0.12, available at: https://cran.r-project.org/web/packages/pheatmap) (Kolde and Kolde, 2015). The correlation map identified two dominant clusters (I and II). The genes attributing to the respective clusters were subsequently analyzed for significantly enriched pathways between HMUg and LMUg predisposition in IPA (adjusted $p < 0.05$).

Results

The microbial community in rumen fluids comprised 7,697 OTUs after filtering, which attributed to 24 phyla, 45 orders, 101 classes, 168 families, and 350 genera across all 20 cows. The filtered transcriptome dataset constituted 12,767 genes.

Microbial analyses

The microbial composition ranked at inverse Simpson values of 87.51 ± 39.13 (mean ± SD) in HMUg cows and 80.95 ± 34.24 (mean ± SD) in LMUg cows, indicating no significant differences between the predispositions (Figure 1A). With lower dietary CP level, alpha diversity numerically decreased, but not significantly, neither in HMUg nor in LMUg cows (Figures 1B,C).

The DESeq2 analysis at genus level revealed 10 DAG between HMUg and LMUg cows, which covered relative abundances between 0.0016 and 2.3668% across all cows (Table 1). Notably, the six most abundant genera (*Prevotella*, *Rikenellaceae_RC9_gut_group*, *Methanobrevibacter*, *Christensenellaceae_R-7_group*, *F082_ge*, and *NK4A214_group*) captured together 51.6% of total rumen abundances and did not differ significantly between HMUg and LMUg cows. Within the DAG, *Unclassified Ruminococcae* exhibited the broadest relative abundance (2.3668%) and manifested a 1.39-fold higher presence in LMUg compared to HMUg animals, whereas *Butyrivibrio* was significantly more abundant in HMUg compared to LMUg cows. Moreover, *Lachnospiraceae_UCG-010* showed significantly higher occurrence in HMUg compared to LMUg, while *Succinivibrionaceae_UCG-002* displayed a 2.24-fold higher occurrence in LMUg than in HMUg animals. However, highest fold changes were observed for *Acetobacter* and *Monoglobus*, with *Acetobacter* depicting 3.25-fold more abundance in HMUg and *Monoglobus* being 2.33-times more prevalent in LMUg animals.

For the comparison of NP and LP diets, 18 DAG were identified in HMUg and 19 DAG in LMUg cows, including four DAG that were previously identified in HMUg–LMUg comparison (Table 2). The reduction of dietary CP content was generally accompanied by high fold changes of genera abundances. For instance, *MVP-15_ge* abundances decreased 133-fold in HMUg, whereas *Sharpea* abundances increased 24-fold in HMUg and 166-fold in LMUg with LP feeding. *Incertae sedis* was less abundant in both, HMUg and LMUg animals, when cows were exposed to the LP diet. Furthermore, the abundances of *Shuttleworthia* increased in HMUg, but decreased in LMUg cows with the LP diet. Following dietary CP reduction, HMUg animals exhibited a further increase of *Acidaminococcus* abundances and of three genera belonging to phylum *Desulfobacterota* (*Desulfovibrio*, *Desulfobulbus*, *Desulfobulbaceae_unclassified*) with mean relative abundances between 0.0015 and 0.0993%. In LMUg animals, the significant increases of *Streptococcus* and *Lachnospiraceae_NK3A20_group* abundances were observed on the LP compared to NP ration. On the contrary, *Alphaproteobacteria_unclassified* and *Succinivibrionaceae_unclassified* abundances decreased remarkably under conditions of LP feeding in LMUg cows at 4.05- and 13.3-times, respectively. Interestingly, *Acetobacter*, whose abundance was significantly lower in LMUg than in

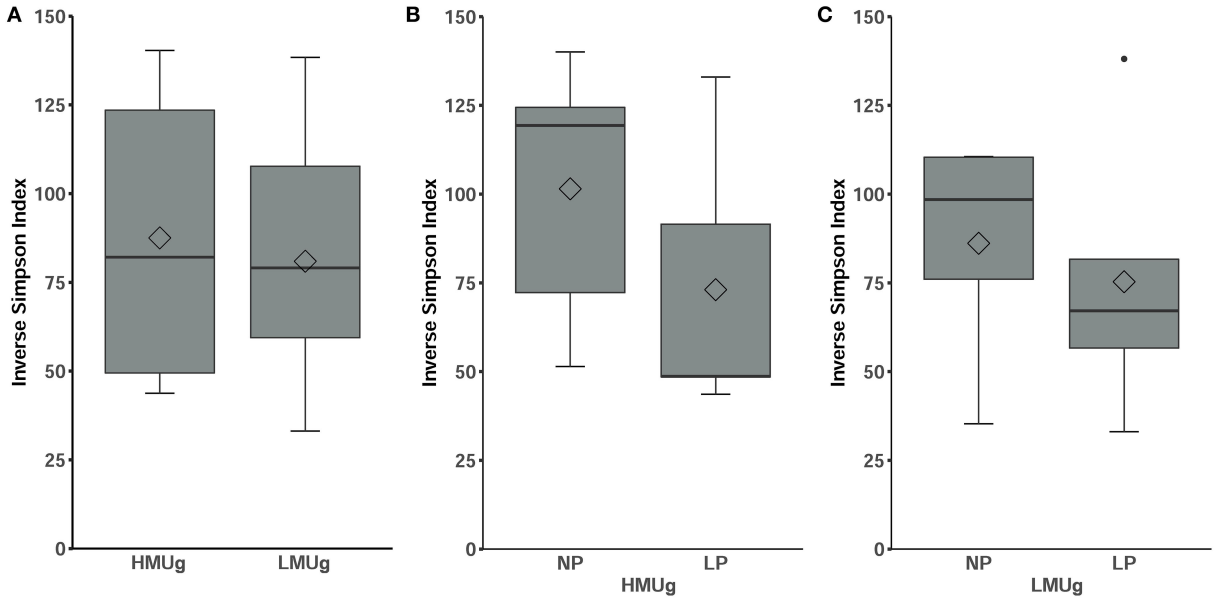

FIGURE 1
Alpha diversity of the microbial community in rumen fluids of **(A)** HMUg and LMUg cows and after lowering dietary N in HMUg **(B)** and LMUg **(C)** cows, respectively. Groups are displayed on the *x*-axis; inverse Simpson values scale the *y*-axis.

TABLE 1 Significantly differentially abundant genera (DAG) in the rumen of HMUg and LMUg cows.

Genus	Relative abundance[a] (%)		Fold change[b]	*p*-value[c]
	HMUg	LMUg		
*Acetobacter**	0.0021	0.0010	3.25	0.0450
Acholeplasmataceae_unclassified	0.0036	0.0075	−2.02	0.0163
Butyrivibrio	0.3949	0.2594	1.37	0.0206
CAG-352	0.3157	0.4505	−1.95	0.0272
Catenibacterium	0.0015	0.0026	−1.71	0.0353
*Lachnospiraceae_UCG-010**	0.0105	0.0060	1.87	0.0418
Monoglobus	0.0121	0.0207	−2.33	0.0021
Ruminococcaceae_unclassified	1.9833	2.7502	−1.39	0.0440
Subdoligranulum	0.0139	0.0195	−1.42	0.0471
Succinivibrionaceae_UCG-002	0.4456	0.8744	−2.24	0.0336

[a]Calculated as mean of the respective group.
[b]Fold changes derive from DESeq2 analysis including trial as fixed effect in the model and refer to HMUg compared to LMUg.
[c]*p*-values < 0.05 indicating significance.
*Displays genera that were also identified as DAG in comparison between NP and LP diets within HMUg and LMUg groups (Table 2).

HMUg cows (Table 1), showed a further 21-fold decrease in LMUg cows upon the reduction of the dietary CP content (Table 2).

Transcriptome analyses

In the filtered transcriptome dataset, 28 genes were identified as differentially expressed in the rumen between HMUg and LMUg cows (Supplementary material 1, p. 1). The highest fold changes were observed for Matrix Metallopeptidase 3 (*MMP3*) and Transition Protein 2 (*TNP2*), with higher transcript abundances in HMUg and LMUg, respectively. Furthermore, 166 genes passed the relaxed significance threshold ($p < 0.01$) for enrichment analysis (Supplementary material 1, p. 2) and revealed 13 significantly enriched pathways that contributed almost exclusively (9 out of 13) to the immune system (Table 3). Antigen Presentation Pathway revealed the

TABLE 2 Significantly differentially abundant genera (DAG) in the rumen comparing normal crude protein (NP) and low crude protein (LP) diet offered to HMUg and LMUg cows.

Genus	Relative abundance[a] (%)		Fold change[b]	*p*-value[c]
	NP	LP		
HMUg				
Acidaminococcus	0.0026	0.0282	−11.52	0.0008
Bacilli_unclassified	0.1169	0.0568	1.80	0.0123
Cardiobacteriaceae_unclassified	0.0217	0.0586	−3.08	0.0361
Desulfobulbaceae_unclassified	0.0007	0.0029	−4.43	0.0056
Desulfobulbus	0.0445	0.1101	−2.71	0.0128
Desulfovibrio	0.0742	0.1605	−2.55	0.0087
horsej-a03	0.0148	0.0015	9.57	0.0395
Incertae_Sedis	0.0144	0.0017	6.54	0.0005
*Lachnospira**	0.0076	0.0013	4.77	0.0049
Lachnospiraceae_UCG-009	0.0049	0.0008	5.31	0.0051
M2PT2-76_termite_group	0.0339	0.0035	7.57	0.0316
MVP-15_ge	0.0136	0.0001	132.57	0.0001
Prevotellaceae_Ga6A1_group	0.0757	0.0118	5.36	0.0073
Pseudobutyrivibrio	0.2100	0.0293	6.22	0.0368
*Sharpea**	0.0015	0.0369	−23.72	0.0055
*Shuttleworthia**	0.0178	0.1586	−10.20	0.0001
Syntrophococcus	0.0855	0.1888	−2.38	0.0411
vadinBE97_ge	0.1724	0.0333	4.71	0.0288
LMUg				
*Acetobacter**	0.0019	0.0001	21.31	0.0068
Alphaproteobacteria_unclassified	0.0158	0.0037	4.05	0.0374
Anaerolineae_unclassified	0.0654	0.3116	−6.11	0.0028
Anaerovibrio	0.0153	0.0048	2.93	0.0085
COB_P4-1_termite_group_ge	0.0096	0.0004	18.24	0.0002
Hungateiclostridiaceae_unclassified	0.0092	0.0222	−2.64	0.0244
*Incertae_Sedis**	0.0139	0.0016	7.35	0.0002
*Lachnospira**	0.0079	0.0020	3.43	0.0250
Lachnospiraceae_NK3A20_group	1.1391	2.8116	−2.88	0.0264
*Lachnospiraceae_UCG-010**	0.0100	0.0020	4.31	0.0078
Pirellulaceae_unclassified	0.0043	0.0131	−3.74	0.0048
Ruminiclostridium	0.0186	0.0476	−2.78	0.0102
*Sharpea**	0.0001	0.0119	−166.22	0.0001
*Shuttleworthia**	0.0600	0.0077	6.45	0.0023
Slackia	0.0009	0.0044	−5.46	0.0472
Streptococcus	0.0057	0.0206	−4.44	0.0071
Succinivibrionaceae_unclassified	0.0092	0.0006	13.29	0.0027
UCG-007 (Oscillospiraceae)	0.0075	0.0008	7.92	0.0154
UCG-012 (Hungateiclostridiaceae)	0.0027	0.0003	7.00	0.0213

[a]Calculated as mean of the respective group.
[b]Fold changes derived from DESeq2 analysis and refer to NP compared to LP.
[c]*p*-values < 0.05 indicating significance.
*Displays genera that were identified significantly differential abundant in more than one contrast.

TABLE 3 Significantly enriched pathways in rumen tissue comparing HMUg and LMUg cows.

Canonical pathways[a]	General function	-log(*p*-value)[b]	Genes
Antigen presentation pathway	IS	7.11	*CD74,BOLA-B,BOLA-DMA,BOLA-DMB,BOLA-DOB,BOLA-DRA,PSMB8,TAPBP*
Interferon signaling	IS	6.19	*IRF1,IRF9,**ISG15**,**MX1**,**OAS1Y**,PSMB8,STAT2*
Complement system	IS	6.19	*C1QA,C1QB,C1R,C1S,C2,CFB,SERPING1*
IL-4 signaling	IS	4.91	*BOLA-B,BOLA-DMA,BOLA-DMB,BOLA-DOB,BOLA-DRA,IL2RG,**IRF4**,JAK3*
B cell development	IS	3.38	*BOLA-B,BOLA-DMA,BOLA-DMB,BOLA-DOB,BOLA-DRA*
Neuroinflammation signaling pathway	IS	2.69	*CXCL12,BOLA-B,BOLA-DMA,BOLA-DMB,BOLA-DOB,BOLA-DRA,IRF7,JAK3,**MMP3**,**PTGS2***
Acetate conversion to Acetyl-CoA	FM/EM	2.17	*ACSS1,ACSS2*
Primary immunodeficiency signaling	IS	1.91	*CD3E,CD4,IL2RG,JAK3*
Glucocorticoid receptor signaling	IS	1.76	*CD3E,BOLA-B,BOLA-DMA,BOLA-DMB,BOLA-DOB,BOLA-DRA,IL2RA,IL2RG,JAK3,**MMP3**,**PTGS2**,**TSC22D3***
Activation of IRF by cytosolic pattern recognition receptors	IS	1.67	*IRF7,IRF9,**ISG15**,STAT2*
Ethanol degradation II	FM/EM	1.63	*ACSS1,ACSS2,AKR1A1*
Glutathione-mediated detoxification	CP	1.62	*GSTA1,**GSTA2**,GSTA4*
Adipogenesis pathway	FM/EM	1.41	*BMP2,CEBPD,FABP4,WNT5A,XBP1*

[a]Terms obtained from Ingenuity Pathway Analysis.
[b]Benjamini–Hochberg adjusted *p*-values <0.05 indicating significance; underlined italic letters display significant inhibition in HMUg (z-score < −2), italic letters indicate predicted inhibition in HMUg (z-score < −2 to 0); bold letters emphasize genes that display fold changes ranking in the top 25 of the filtered transcriptome datasets.
IS, immune system; FM/EM, fat/energy metabolism; CP, cell protection.

most prominent adjusted *p*-value. In addition, Interferon Signaling and Neuroinflammation Signaling pointed as two further candidates, displaying significant inhibition in HMUg. Whereas, Ubiquitin Like Modifier 15 (*ISG15*) and MX Dynamin Like GTPase 1 (*MX1*) were included in the former, *MMP3* and Prostaglandin-Endoperoxide Synthase 2 (*PTGS2*) contributed to the latter. Interestingly, these four genes also ranked under the top 25 genes in terms of absolute fold changes among all annotated genes in the filtered dataset (Supplementary material 1). However, Serine Protease 2 (*PRSS2*) and UL16 Binding Protein 17 (*ULBP17*) displayed the most prominent fold changes.

Considering the reduction of dietary CP in HMUg cows, no DEGs were observed, but 68 genes passed the relaxed significance threshold for pathway analysis. This list culminated in the identification of Intrinsic Prothrombin Activation Pathway and MSP-RON Signaling Pathway as significantly enriched (Table 4; Supplementary material 2, p. 1). In contrast, LMUg animals showed 1,220 DEGs and 139 significantly enriched pathways when dietary CP content was reduced (Table 4; Supplementary material 2, p. 2, 3). Whereas, Oxidative Phosphorylation and Semaphorin Neuronal Repulsive Signaling Pathway were significantly activated when LMUg cows were fed the LP ration, additional 79 pathways were identified to be significantly inhibited, mainly representing pathways of immune response and energy metabolism as well as pathways that contribute to cell cycle and cell maintenance (Table 4).

Integration of microbial and transcriptome data

The multivariate discriminant analysis (sPLS-DA) revealed a microbial subset comprising 50 genera in the first component and 5 genera in the second component separating HMUg and LMUg cows' rumen microbial communities (Figure 2A; Supplementary material 3, p. 1). Two genera contributed to both components, giving a subset of 53 genera that distinguished HMUg and LMUg cows. Eight genera of the subset were also identified as DAG in HMUg–LMUg contrast. The transcriptome subset included 100 genes, comprising 50 genes in each component (Figure 2B; Supplementary material 3, p. 2). Twelve genes thereof were uncovered as DEG between HMUg and LMUg.

TABLE 4 Significantly enriched pathways between normal protein (NP) and low protein (LP) diet in rumen tissue of HMUg and LMUg cows.

	Canonical pathways[a]	General function	*z*-score	-log(adj. *p*-value)[b]
HMUg				
	Intrinsic prothrombin activation pathway	IS		2.9
	MSP-RON signaling pathway	IS		2.75
LMUg				
	Insulin receptor signaling	EM	−2.294	4.49
	PI3K signaling in B lymphocytes	IS	−3.128	4.16
	CNTF signaling	IS	−2.673	4.07
	RAN signaling	CC	−2.828	4.07
	Virus entry *via* endocytic pathways	IS		3.91
	Protein ubiquitination pathway	CC		3.82
	GM-CSF signaling	IS	−2.138	3.82
	UVA-induced MAPK signaling	CC	−2.887	3.82
	p70S6K signaling	CF	−2.982	3.64
	Kinetochore metaphase signaling pathway	CC	−1.698	3.47
	Signaling by Rho family GTPases	CM	−1.8	3.37
	IL-3 signaling	IS	−2.324	3.36
	Aldosterone signaling in epithelial cells	CS	−1.941	3.33
	Ephrin A signaling	CS		3.22
	IGF-1 signaling	EM	−2.138	3.18
	*Semaphorin neuronal repulsive signaling pathway**	CM	2.065	2.54
	*Oxidative phosphorylation**	EM	2.84	2.39

[a]Terms obtained from Ingenuity Pathway Analysis.
[b]Benjamini–Hochberg adjusted *p*-values < 0.05 indicating significance; underlined italic letters indicate significant activation (z-score>2) or inhibition (z-score < -2) under NP, italic letters indicate predicted inhibition (z-score < -2 to 0) under NP; for LMUg only the top 15 (adjusted *p*-value) significantly inhibited pathways and the two significantly activated pathways (*indicates activation) under NP are displayed.
IS, immune system; FM/EM, fat/energy metabolism; CP, cell protection; CC, cell cycle; CF, cell function; CM, cell maintenance; CS, cell signaling.

The correlation analysis between the microbial and the gene subset separating HMUg and LMUg cows detected 157 significantly correlated microbial genera–gene pairs. The two pronounced clusters I and II were generated by two groups of microbes (A and B) and two groups of genes (I and II) that correlated contrarily (Figure 3; Supplementary material 3). In detail, Cluster I depicted gene Group I (27 genes) being positively correlated to microbial Group A (24 genera) and negatively correlated to microbial Group B (29 genera). Cluster II on the other hand showed negative correlations between gene Group II (73 genes) and microbial Group A, but positive correlations between gene Group II and microbial Group B.

The subsequent pathway analysis of gene Groups I and II revealed significant enrichment of Urea Cycle, Arginine Degradation and Citrulline Metabolism for Group I, which was mainly determined by Arginase 1 (*ARG1*) expression (Table 5). The *ARG1* expression was negatively correlated with *UCG-005 abundances* ($r = -0.64$, $p < 0.01$; Figure 3; Supplementary material 4). Furthermore, the differential expression of Glutamine-fructose-6-phosphate transaminase 1 (*GFPT1*) in Group I contributed to the significant enrichment of UDP-N-acetyl-D-glucosamine biosynthesis II (Table 5). Moreover, *GFPT1* was negatively correlated with *UCG-010_ge* ($r = -0.68$, $p < 0.01$), *Lachnospiraceae_UCG-006* ($r = -0.56$, $p = 0.011$), and *Roseburia* abundances ($r = -0.56$, $p = 0.01$) (Figure 3; Supplementary material 4). A higher expression level of *GFTP1* was accompanied by increased abundances of *Lachnospiraceae_UCG-010* ($r = 0.52$, $p = 0.02$), *Acetobacter* ($r = 0.56$, $p = 0.01$), and *Shuttleworthia* ($r = 0.55$, $p = 0.01$), which had been pre-identified as DAG in the comparison of HMUg and LMUg cows (Figure 3; Table 1; Supplementary material 4).

Gene Group II revealed significant enrichment of four pathways that correspond to the immune system, including B Cell Development, IL-4 Signaling and Antigen Presentation Pathway, in overlap to the analysis of ruminal DEGs (Tables 3, 5). Particularly, genes of the Major Histocompatibility Complex (*BOLA-B*, *BOLA-DOB*, and *BOLA-DRA*) as well as Interferon Regulatory Factor 4 (*IRF4*) contributed to the significant enrichment of immune pathways in Group II (Table 5). *BOLA-DRA* transcripts were significantly less abundant in HMUg compared to LMUg animals (DEG) and revealed positive

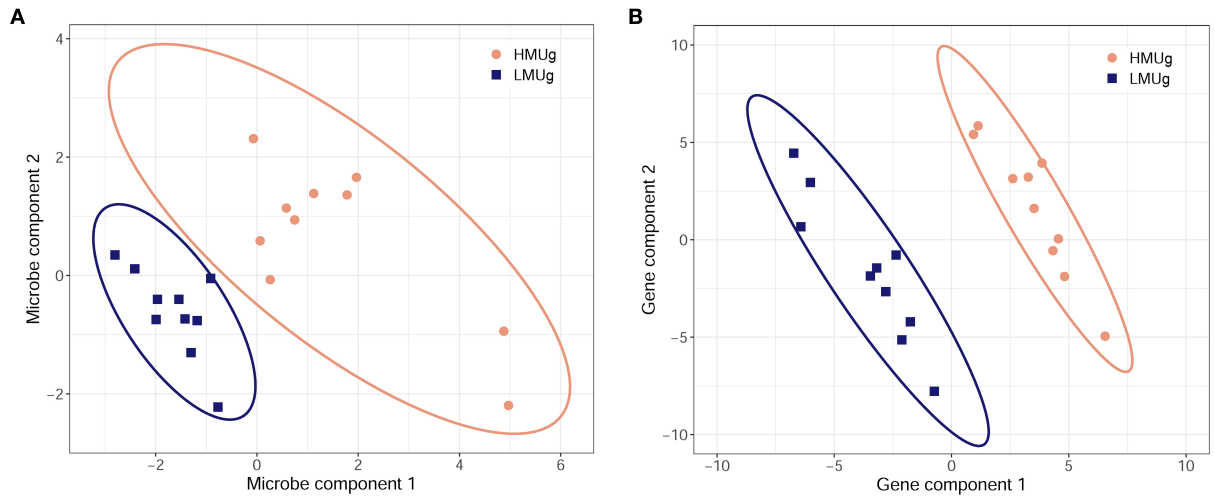

FIGURE 2
Sparse partial least square-discriminant analysis of HMUg and LMUg cows, driven by **(A)** 53 key microbial genera in the rumen fluid and by **(B)** 100 key genes in the rumen tissue.

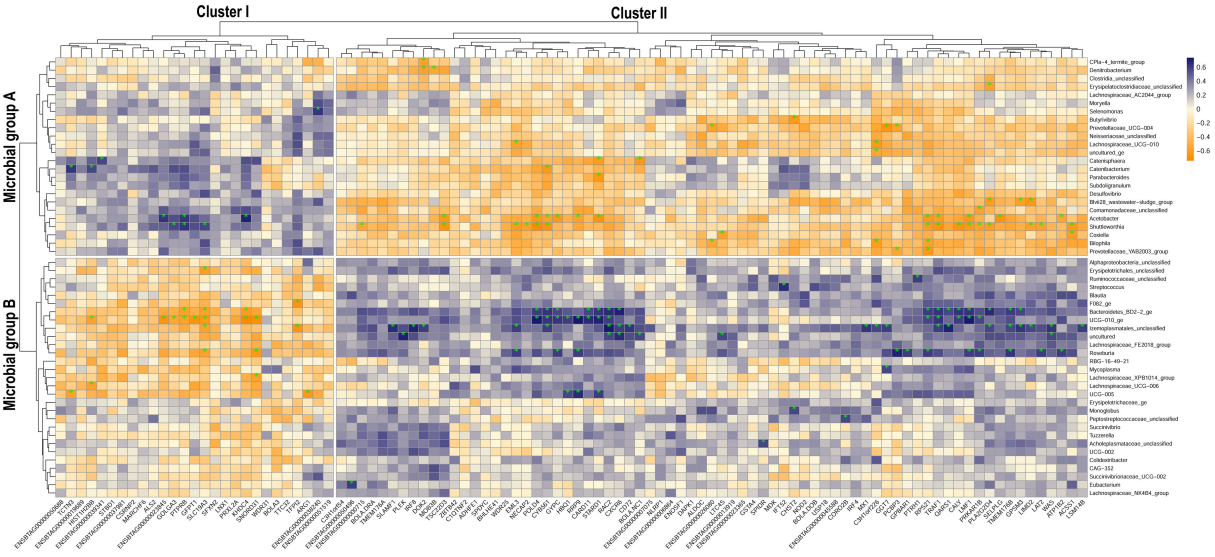

FIGURE 3
Correlation map of ruminal host-microbiota interaction depicting the relationship of microbial subset abundances and gene subset transcripts that separate HMUg and LMUg rumen profiles. Subsets derived from sPLS-DA analysis; estimated correlations are based on the microbial abundances and gene transcripts in HMUg and LMUg cows; *indicates significant correlation ($p < 0.01$).

correlations with abundances of *Roseburia* ($r = 0.53$, $p = 0.02$) and *Lachnospiraceae_FE2018_group* ($r = 0.52$, $p = 0.02$) (Figure 3; Supplementary materials 1, 4). In addition, the presence of Gamma-Glutamyltransferase 7 (*GGT7*) and Leukotriene C4 Synthase (*LTC4S)* in Group II caused the significant enrichment of Leukotriene Biosynthesis Pathway (Table 5). The transcript abundance of both genes was highly correlated with the abundance of *Prevotellaceae_UCG-004* (*GGT7*: $r = -0.67$, $p < 0.01$, *LTC4S*: $r = -0.52$, $p = 0.02$) (Figure 3; Supplementary material 4). Moreover, *GGT7* transcripts showed further a prominent negative correlation to *Lachnospiraceae_UCG-010* abundances ($r = -0.05$, $p = 0.02$), whereas *LTC4S* transcripts negatively correlated with the occurrence of *Coxiella* ($r = -0.58$, $p < 0.01$).

Furthermore, *UCG-010_ge* genus attracted interest due to its particularly high correlations with transcript abundances

TABLE 5 Significantly enriched pathways between HMUg and LMUg cows based on gene Groups I and II derived from ruminal host-microbiota interaction analysis (Figure 3).

Ingenuity canonical pathways[a]	–log(*p*-value)[b]	Genes	Ratio[c]	Cluster
Urea cycle	1.85	*ARG1*	0.17	Group I
Arginine degradation VI (Arginase 2 Pathway)	1.85	*ARG1*	0.17	Group I
UDP-N-acetyl-D-glucosamine biosynthesis II	1.85	*GFPT1*	0.17	Group I
Arginine degradation I (arginase pathway)	1.85	*ARG1*	0.25	Group I
Citrulline biosynthesis	1.78	*ARG1*	0.11	Group I
Superpathway of citrulline metabolism	1.63	*ARG1*	0.07	Group I
B cell development	1.87	*BOLA-B,BOLA-DOB,BOLA-DRA*	0.07	Group II
IL-4 signaling	1.87	*BOLA-B,BOLA-DOB,BOLA-DRA,IRF4*	0.043	Group II
Antigen presentation pathway	1.87	*BOLA-B,BOLA-DOB,BOLA-DRA*	0.08	Group II
Leukotriene biosynthesis	1.6	*GGT7,LTC4S*	0.14	Group II

[a]Terms obtained from Ingenuity Pathway Analysis (IPA).
[b]Benjamini–Hochberg adjusted *p*-values < 0.05 indicating significance.
[c]contribution of the gene dataset (Gene Groups I or II) to respective pathway enrichment, calculated by IPA.

of KH Domain Containing 4 (*KHDC4*; $r = -0.72$, $p = 0.0003$), Ribosomal RNA Processing 9 (*RRP9*; $r = 0.71$, $p < 0.0005$), Lipase Maturation Factor 1 (*LMF1*; $r = 0.73$, $p < 0.0003$) and DNA Polymerase Delta 4 (*POLD4*; $r = 0.73$, $p = 0.0002$). Furthermore, *POLD4* expression showed a strikingly high negative correlation with the abundance of *Shuttleworthia* ($r = -0.74$, $p = 0.00017$), whereas *KHDC4* transcripts showed a prominent correlation with *Acetobacter* occurrences ($r = 0.69$, $p = 0.0008$).

Discussion

In regard to the underlying biological principles of N-metabolism in ruminants, it was hypothesized that the predisposed variance in MU concentration is—in complement to inherent post-ruminal processes—mainly attributed to the ruminal mechanisms that differentiate between HMUg and LMUg cows and thus lead to N-utilization and N-excretion variance in genetically divergent cow cohorts. Specifically, the rumen microbial N-utilization, the absorption and diffusion processes of the rumen epithelium and the ruminal interaction between microbes and host epithelium were thought to characterize the ruminal make up of HMUg and LMUg cows.

Microbial analyses

Although no differences in the overall microbial diversity have been uncovered between HMUg and LMUg cows, some DAGs were identified. These DAG might (i) influence the amount of ruminal NPN in forms of NH_3 or NH_4^+, (ii) determine the utilization of blood urea by ureolytic activity, or (iii) influence the permeability for the transport of NH_3, NH_4^+ and urea across the rumen epithelium.

Specifically, LMUg animals displayed a significantly higher abundance of *Succinivibrionaceae_UCG-002*, which is an ureolytic genus in the rumen. Ureolytic bacteria metabolize urea to CO_2 and NH_3. The latter can be further utilized by other ruminal genera for microbial growth, which subsequently reduces the ruminal NPN-pool (Rosendahl, 2015). Moreover, the conversion of urea into NH_3 increases the urea concentration gradient between blood and the rumen lumen. This massively alters the urea transport rate, reduces blood urea concentration and promotes N-fixation into microbial protein (Rosendahl, 2015; Jin et al., 2016). Interestingly, significantly lower blood urea concentrations and lower blood urea pool sizes were observed in cows displaying the LMU compared to the HMU phenotype in a recent study (Müller et al., 2021). Furthermore, unclassified genera of the *Succinivibrionaceae* family have already been suspected as being more abundant in ruminants with increased N-use efficiency (Jin et al., 2016).

LMUg animals additionally exhibited significantly higher abundances of *unclassified Ruminococcaceae*. Various *Ruminococcaceae* genera correlated with the N-metabolism in goats, including *Ruminococcus_2* displaying significantly higher abundances in animals with higher N utilization efficiency (Wang et al., 2019). Moreover, the *Ruminococcaceae* family has been suggested to play a fundamental role in amino acid and protein metabolism in the bovine rumen, and the *Ruminococcaceae NK4A214 group* has been attributed a role to N-recycling in cattle (Pacífico et al., 2021). Hu et al. (2019) found a positive correlation between the *Ruminococcus* abundances and the gap junctions in the rumen epithelium of yaks and postulated an absorptive response of the rumen epithelium to *Ruminococcus* metabolites. Some *Ruminococcaceae* genera were also reported to have urease activity (Patra and Aschenbach, 2018), and thus it might be conceivable that *unclassified Ruminococcaceae* increases the urea concentration gradient

across the rumen wall and support the diffusion of urea in LMUg animals.

Furthermore, HMUg animals showed higher abundances of *Butyvibrio*, which is one of the most relevant butyrate producers in the rumen (Meehan and Beiko, 2014; Henderson et al., 2015; Kong et al., 2020). Interestingly, two *Butyvibrio* species, namely, *B. fibrisolvens* and *B. proteoclasticus*, were postulated as key species involved in protein degradation in the bovine rumen (Wallace and Brammall, 1985; Attwood and Reilly, 1995; Henderson et al., 2015). Due to their high proteolytic activities, they belong to the hyper-ammonia producing bacteria species (HAB) (Hartinger et al., 2018). Also, HAB species are known to negatively influence the N-efficiency of ruminants, since they cause steep increases of NH_3, which cannot be directly utilized by the rumen microbes and thus leads to enhanced NH_3 efflux into the bloodstream (Bento et al., 2015; Hartinger et al., 2018). Based on phenotypic categorization into HMU and LMU cow groups, HMU cows were indeed shown to exhibit numerically higher ruminal NH_3 concentrations than LMU cows (Müller et al., 2021). Furthermore, higher *Butyrivibrio* abundances were found to be correlated with lower N-recycling efficiency in beef cattle (Alves et al., 2020).

The influence of different CP levels on the rumen microbial community has been broadly evidenced (Lapierre et al., 2005; Aguiar et al., 2014; Patra and Aschenbach, 2018; Müller et al., 2021). In general, it is known that lower CP levels in the diets of dairy cows reduce rumen fluid NH_3 and NH_4^+, enhance the relative blood urea reflux and faster N-assimilation and N-fixation into microbial protein.

In both HMUg and LMUg animals, the reduction of dietary CP content induced significant changes in microbial abundances, which might be caused by microbial N-use competition and the general dietary nutrient availability. In this context, *Sharpea* abundances increased simultaneously with the decline of *Incertae_sedis* abundances in both, HMUg and LMUg cows, when feed-CP content was lowered. *Sharpea* is known to digest short-chain carbohydrates (Dias et al., 2017; Trabi et al., 2020), while *Incertae_sedis* (aggregated genus of the *Ethanoligenenaceae* and *Ruminococcaceae* families) primarily converts cellulose (Suen et al., 2011; Bailoni et al., 2021). The metabolic processes and the resulting microbial growth of cellulose digesters are generally slower than the growth of microorganisms, which metabolize the short-chain sugars. However, both microorganisms need to assimilate N components for their microbial growth (Russell et al., 2009). It might be conceivable that *Sharpea* depressed the growth of *Incertae_sedis* due to comparatively faster N-assimilation, when CP was lowered. *Shuttleworthia* attracted further interest since its abundance increased in HMUg, whereas significant abundance decreases were identified in the LMUg group. Several studies reported feed induced abundance alterations of *Shuttleworthia*, which were related to the dietary starch–fiber ratio and the energy level of the diet (Plaizier et al., 2017; Kotz et al., 2020; Tun et al., 2020). Furthermore, Zhang et al. (2019) postulated significantly different abundances between high and low yielding dairy breeds and identified *Shuttleworthia* as a genus displaying high relative abundances, which is in common with the present study. These findings promote *Shuttleworthia* as highly vulnerable and strongly adaptable genus, which might be of importance in characterizing the different ruminal adaption mechanisms in LMUg and HMUg on low CP levels.

Moreover, the abundances of *Acidaminococcus* and three genera of *Desulfobacterota* phylum (*Desulfovibrio*, *Desulfobulbus*, and *Desulfobulbaceae_unclassified*) increased in HMUg cows with the reduction of dietary CP content. Interestingly, several species of these genera have been previously categorized as HAB (Attwood and Reilly, 1995; Eschenlauer et al., 2002; Loubinoux et al., 2002). The growth of HAB is known to be suppressed by high NH_3 and NH_4^+ concentrations in a self-regulating manner (Sales et al., 2000). Interestingly, phenotypic HMU animals displayed significant decreases in previously high ruminal NH_3 and NH_4^+ concentrations following CP reduction, whereas in LMU phenotypes only moderate ruminal NH_3 and NH_4^+ declines were observed (Müller et al., 2021). Since genera abundances of the *Desulfobacterota* phylum appeared to be significantly increased exclusively in HMUg, it is conceivable that their growth was inhibited by high ruminal NH_3 and NH_4^+ levels in HMUg under NP diet feeding, and that the reduction of dietary N might have had a positive effect on their growth.

In LMUg cows, dietary CP reduction led to increased abundances of *Streptococcus*, which is one of the most important proteolytic members in the rumen fluid (Attwood and Reilly, 1995; Ling and Armstead, 1995). *Streptococcus* utilizes rapidly assimilable energy sources (starch), is known to possess ureolytic activity and is highly adaptable to various N-sources, which could provide a competitive advantage under N scarcity (Jin et al., 2016; Hartinger et al., 2018; Patra and Aschenbach, 2018).

Under conditions of LP diet feeding, LMUg cows further exhibited a significant abundance decline of *Acetobacter*, which had already been identified as significantly less abundant in LMUg compared to HMUg cows in this study (contrast HMUg–LMUg). The further massive abundance decline in LMUg cows from NP to LP feeding suggests *Acetobacter* an adaptable capacity to different N levels in the rumen, which may contribute to differential N-utilization in LMUg cows.

Transcriptome

While the metabolic activity of the microbial community substantially affects the ruminal protein-N to NPN ratio, the gene expression in the rumen villi impact the ruminal milieu (e.g., active and passive transport processes) and thus the composition of the microbial community (e.g., immune

modulation) (Penner et al., 2011; Steele et al., 2016). The pathway analysis of DEGs in rumen villi revealed the enrichment of immune response pathways, which were activated in LMUg compared to HMUg cows. Interestingly, it was recently reported in humans that the intestinal mucosa was able to respond to the microbial products with the targeted expression of receptors, immune genes and signaling molecules to maintain intestinal homeostasis (Tannock and Liu, 2020). In addition, the concept of "immune selection" was described as a way of maintaining host–microbe relationships in the long term as a result of host immune activity (Peterson et al., 2007). In this context, Bessman and Sonnenberg (2016) elucidated the essential role of Antigen Presentation and the Major Histocompatibility Complex, which dictated the host T-cell reactivity to biochemical products of commensal bacteria by the expression of innate lymphoid cells. Notably, the highest *Z*-score with significant activation in LMUg compared to HMUg cows was found for the Antigen Presentation Pathway in this study. Furthermore, the pathways of IL-4 Signaling, B Cell Development and Glucocorticoid Receptor identified in LMUg–HMUg contrast were enhanced, predominantly by Bovine Leukocyte Antigens (BOLA), which are prominent drivers of the bovine Major Histocompatibility Complex (Ellis and Ballingall, 1999; Lewin et al., 1999). Moreover, Zhang et al. (2021) reported a simultaneous slight activation of the Antigen Presentation Pathway with a significantly higher expression of *PRRS2* in the rumen tissue of calves with comparatively higher N-efficiency. In fact, *PRSS2* showed a severely higher expression in LMUg compared to HMUg cows. Thus, both antigen presentation and *PRRS2* might constitute key candidates distinguishing ruminal N-utilization of LMUg and HMUg. Also, *PRSS2* has further been associated with feed efficiency and performance traits in cattle (Abo-Ismail et al., 2013), which might be of interest since MX Dynamin Like GTPase 1 (*MX1*) and ISG15 Ubiquitin Like Modifier (*ISG15*) were similarly found to be higher expressed in more feed-efficient heifers, which simultaneously possessed more active immune systems (Paradis et al., 2015). All of these three genes displayed remarkably higher expression levels in LMUg compared to HMUg cows in this study driving the enrichment of Interferon and Neuroinflammation Signaling.

Among the significantly DEG, Matrix Metallopeptidase 3 (*MMP3)* and *TNP2* exhibited the broadest expression differences between cow groups. Metalloproteases were inferred to influence the destruction of tissues by remodeling the extracellular matrix (Coussens et al., 2003; Abendaño et al., 2013) and *MMP3* in specific has been recognized to influence the remodeling of adipose tissue in cattle (Jeong et al., 2017). It might be conceivable that *MMP3* also attributes to structure processes in the rumen tissue and thus contribute to different absorption and diffusion capacities between HMUg and LMUg epithelia, regarding ruminal NH_3, NH_4^+, and urea fluxes. Furthermore, *TNP2* has been reported in the context of male fertility (Pasquariello, 2015), but up to now, its function in the rumen of cattle as well as its possible role in N-metabolism remains uncertain. However, given the relationship of fertility traits and immune system activities in dairy cows, the differential expression of *TNP2* demands further investigation whether LMUg and HMUg cows differ with respect to fertility parameters (Chebel et al., 2004; Hansen et al., 2004; Moore et al., 2005; König et al., 2008; Weiss et al., 2009; Hurley, 2014).

Although the reduction in dietary CP content did not trigger significantly differential gene expression in rumen tissue of HMUg animals, LMUg cows exhibited severe adaptations in their transcriptome, as characterized by more than 1,000 DEG and 139 significantly enriched pathways. The pathways identified contributed almost exclusively to the immune system, the energy metabolism and the cell cycle. Considering the significantly higher plasma urea pool size during NP compared to LP feeding in cows phenotyped for LMU (Müller et al., 2021), and the general role of immune cells in the rumen wall in protecting the host against infiltration of toxic compounds, the enrichment of these pathways may indicate the necessity for the rumen epithelia to balance the N-gradient and maintain immune homeostasis (Bessman and Sonnenberg, 2016; Tannock and Liu, 2020; Müller et al., 2021). In regard to the severe energy requirements of the immune system (Gleeson et al., 2004; Kvidera et al., 2017), it might further be speculated that the observed effects on energy metabolism pathways in LMUg cows fed the LP ration result from an altered energy demand to stabilize the ruminal N balance.

Relationship between microbial genera and transcriptome subsets

The correlation analysis between microbial and host gene subsets reflecting the rumen microbiome—host interaction, which distinguished the rumen profiles of HMUg and LMUg cows, identified two main clusters. The pathway analysis of Gene Group I revealed the enrichment of Urea Cycle and Arginine Metabolism, primarily due to increased expression of Arginase 1 (*ARG1*) in LMUg compared to HMUg cows. Arginase is an enzyme that catalyzes the last reaction step in the urea cycle by facilitating the synthesis of arginine into ornithine and urea. In general, the urea cycle is performed predominantly by hepatocytes in the liver (Emmanuel, 1980; Razmi et al., 2005). However, Beck et al. (2009) stated the expression of *ARG1* also in the rumen epithelium of cattle, which implies catabolism of N-metabolites *via* the urea cycle in the rumen wall. Similarly, Emmanuel (1980) identified low *ARG1* expression in the rumen epithelium of sheep. Considering the conversion of NH_3 into urea, the prominent negative correlation between *ARG1* transcripts and *UCG-005* abundances calls for further research. Although *UCG-005* has not been mentioned directly in the context of N-efficiency in ruminants yet, Amat et al. (2021)

predicted *UCG-005* a negative impact on *Methanobrevibacter* abundances in yearling heifers, which is known to contain various N-fixating strains (Poehlein et al., 2018).

Further insights into a possible differentiation of HMUg and LMUg were displayed by Gene Group II, which revealed the enrichment of four immune system pathways, particularly attributed to the expression of *BOLA* genes. Interestingly, *BOLA* polymorphisms were not only reported to determine inherent and acquired immune defenses (Rupp and Boichard, 2003), they were also capable to modulate the composition of colostrum microbiota in dairy cows (Derakhshani et al., 2018). Furthermore, *BOLA* gene expression is known to influence the symbiotic microbial community in the gastrointestinal tract of humans (De Palma et al., 2010), mice (Toivanen et al., 2001), and fish (Bolnick et al., 2014). In the present study, *BOLA-DRA* transcripts were significantly more abundant in LMUg compared to HMUg cows and showed further striking positive correlations to *Roseburia* abundances. The contribution of *Roseburia* to immune responses and to the N-metabolism in cattle, goats, and rabbits has been postulated by various authors (Wang et al., 2011; Alves et al., 2020; Sun et al., 2020). The observed higher abundances of *Roseburia* and *BOLA-DRA* transcripts in LMUg compared to HMUg cows might picture a microbial–host interplay that contributes to the distinction of ruminal patterns between HMUg and LMUg cows by affecting immune response and N-metabolism-associated processes.

In addition to the correlation with *Roseburia*, *BOLA-DRA* transcripts revealed a severe positive correlation to *Lachnospiraceae_FE2018_group* abundances, which belongs—like as *Roseburia*—to the *Lachnospiraceae* family. *Lachnospiraceae_FE2018_group* had been mentioned in terms of intestinal barrier function, antioxidant balance, and intestinal inflammation in the gut immune system of broilers (Kong et al., 2020; Liu et al., 2021). Moreover, Wang et al. (2021) postulated a significant positive correlation between *Lachnospiraceae family* and intestinal inflammation in dairy cows. The *Lachnospiraceae* family was also found to be significantly more abundant in beef steers with low N-retention efficiency and higher urinary N-excretion (Alves et al., 2020). Furthermore, high abundances of *Lachnospiraceae* were uncovered in the colon and caecum of goats with low N-utilizing phenotype (Wang et al., 2019). Additionally, *Lachnospiraceae* occurrences were found to be correlated with a low feed conversion rate and low feed efficiency in beef steers (Hernandez-Sanabria et al., 2010; Carberry et al., 2012).

Conclusion

This study focused on the ruminal background of Holstein cows with predisposed higher or lower MU concentration (HMUg–LMUg) as well as on their adaptation to low CP diets. Compared to HMUg predisposition, LMUg cows displayed higher occurrences of ureolytic genera, such as *Succinivibrionaceae_UCG-002* and *Ruminococcaceae_unclassified*, which might cause lower blood urea concentrations and lower blood urea pool sizes in LMU phenotypes. Similarly, the higher ruminal NH_3 concentrations in HMU phenotypes might be attributed to the higher occurrences of high ammonia-producing species hosted by HMUg cows. If enhanced immune responses that were uncovered in LMUg compared to HMUg cows' rumen epithelia influence the epithelial barrier for NPN molecules and thus drive MU phenotype distinction can only be speculated at this point. However, the downregulation of immune responses and energy metabolism pathways in LMUg cows fed a low CP diet might have indicated energy wasting efforts of the rumen epithelia to maintain ruminal N-balance when LMUg cows were exposed to diets with normal CP content. Considering the important role of Arginase in the urea cycle, the observed interplay between *ARG1* expression and *UCG-005* occurrences is specifically proposed for future research on microbe–host interactions. The ruminal patterns identified for HMUg and LMUg cows contribute to a deeper insight into MU predisposition of Holsteins and attribute to the optimization of N-utilization in tandem with the reduction of N-emissions on dairy farms by future breeding selection strategies.

Data availability statement

The data presented in the study are deposited in the ArrayExpress Microarray Database at EBI (RNA datasets) and in the BioSample Database (16S Amplicon data) repository, accession numbers E-MTAB-9901 (RNA datasets) and PRJNA856508 (16S Amplicon data).

Ethics statement

The animal study was reviewed and approved by Ethics Committee of the State of Mecklenburg-Western Pomerania (State Office for Agriculture, Food Safety and Fisheries; LALLF M-V7221.3-2-019/19). Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

CM, MP, and BK organized and carried out animal husbandry. HH and HR jointly collected the samples, performed statistical analysis, and interpreted the data. HH conducted laboratory work and wrote the manuscript. NT performed RNA and 16S sequencing. SP supported data analysis. DS assisted in data interpretation. NR, BK, and KW conceptualized and supervised the study. All authors reviewed the final manuscript and provided critical feedback.

Funding

As a part of the BlueCow Project (FKZ: 281B101516) this study has received funding from the Federal Ministry of Food and Agriculture.

Conflict of interest

Author DS was employed by IT-Solutions for Animal Production, Vereinigte Informationssysteme Tierhaltung w.V. (vit).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.939711/full#supplementary-material

SUPPLEMENTARY MATERIAL 1
p. 1: Significantly differentially expressed genes (DEG) between HMUg and LMUg cows. Log2fold changes refer to HMUg. p. 2: Gene set for enrichment analysis (p-values are below 0.01, and the corresponding q-values are below 0.6) between HMUg and LMUg cows. Log2fold changes refer to HMUg.

SUPPLEMENTARY MATERIAL 2
p. 1: Gene set for enrichment analysis (p-values are below 0.01, and the corresponding q-values are below 0.6) between NP and LP diet in HMUg cows. Log2fold changes refer to HMUg. p. 2: Gene set for enrichment analysis (p-values are below 0.01, and the corresponding q-values are below 0.6) between NP and LP diet in LMUg cows. Log2fold changes refer to NP. p. 3: Significantly enriched pathways (adjusted p-values are below 0.05) between NP and LP diet in LMUg cows.

SUPPLEMENTARY MATERIAL 3
p. 1: Microbial subset (53 genera) generated by sparse Partial Least Squares Discriminant Analysis (sPLS-DA) separating HMUg and LMUg cows. p. 2: Gene subset (100 genes) generated by Partial Least Squares Discriminant Analysis (sPLS-DA) separating HMUg and LMUg cows.

SUPPLEMENTARY MATERIAL 4
p. 1: The p-values of the correlation analysis between microbial genera abundances and gene transcripts separating HMUg and LMUg cows (Figure 3). p. 2: Correlation coefficients (r-values) of the correlation analysis between microbial genera abundances and gene transcripts separating HMUg and LMUg cows (Figure 3).

References

Abdoun, K., Stumpff, F., and Martens, H. (2006). Ammonia and urea transport across the rumen epithelium: a review. *Anim. Health Res. Rev.* 7, 43–59. doi: 10.1017/S1466252307001156

Abendaño, N., Juste, R. A., and Alonso-Hearn, M. (2013). Anti-inflammatory and antiapoptotic responses to infection: a common denominator of human and bovine macrophages infected with Mycobacterium avium subsp. paratuberculosis. *BioMed. Res. Int.* 2013:7. doi: 10.1155/2013/908348

Abo-Ismail, M., Kelly, M., Squires, E., Swanson, K., Bauck, S., and Miller, S. (2013). Identification of single nucleotide polymorphisms in genes involved in digestive and metabolic processes associated with feed efficiency and performance traits in beef cattle. *J. Animal Sci.* 91, 2512–2529. doi: 10.2527/jas.2012-5756

Aguiar, S. C. D., Paula, E. M. D., Yoshimura, E. H., Santos, W. B. R. D., Machado, E., Valero, M. V., et al. (2014). Effects of phenolic compounds in propolis on digestive and ruminal parameters in dairy cows. *Rev. Brasil. Zootecn.* 43, 197–206. doi: 10.1590/S1516-35982014000400006

Alves, K. L. G. C., Granja-Salcedo, Y. T., Messana, J. D., De Souza, V. C., Ganga, M. J. G., Colovate, P. H. D., et al. (2020). Rumen bacterial diversity in relation to nitrogen retention in beef cattle. *Anaerobe* 67, 102316. doi: 10.1016/j.anaerobe.2020.102316

Amat, S., Holman, D. B., Schmidt, K., Menezes, A. C. B., Baumgaertner, F., Winders, T., et al. (2021). *The nasopharyngeal,* ruminal, and vaginal microbiota and the core taxa shared across these microbiomes in virgin yearling heifers exposed to divergent in utero nutrition during their first trimester of gestation and in pregnant beef heifers in response to mineral supplementation. *bioRxiv*. 9, 2011. doi: 10.3390/microorganisms9102011

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Attwood, G., and Reilly, K. (1995). Identification of proteolytic rumen bacteria isolated from New Zealand cattle. *J. Appl. Bacteriol.* 79, 22–29. doi: 10.1111/j.1365-2672.1995.tb03119.x

Bailoni, L., Carraro, L., Cardin, M., and Cardazzo, B. (2021). Active rumen bacterial and protozoal communities revealed by RNA-based amplicon sequencing on dairy cows fed different diets at three physiological stages. *Microorganisms* 9, 754. doi: 10.3390/microorganisms9040754

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

Beck, D., Jiang, H., and Zhang, L. (2009). Elucidating the evolutionary relationships among bos taurus digestive organs using unigene expression data. *Int. J. Evolut. Biol.* 2010, 803142. doi: 10.4061/2009/803142

Bento, C. B. P., De Azevedo, A. C., Detmann, E., and Mantovani, H. C. (2015). Biochemical and genetic diversity of carbohydrate-fermenting and obligate amino acid-fermenting hyper-ammonia-producing bacteria from Nellore steers fed tropical forages and supplemented with casein. *BMC Microbiol.* 15, 28. doi: 10.1186/s12866-015-0369-9

Bergen, W. G. (2021). "Amino acids in beef cattle nutrition and production," in *Amino Acids in Nutrition and Health* (New York, NY: Springer), 29–42. doi: 10.1007/978-3-030-54462-1_3

Bessman, N. J., and Sonnenberg, G. F. (2016). Emerging roles for antigen presentation in establishing host–microbiome symbiosis. *Immunol. Rev.* 272, 139–150. doi: 10.1111/imr.12425

Bolnick, D. I., Snowberg, L. K., Caporaso, J. G., Lauber, C., Knight, R., and Stutz, W. E. (2014). Major H istocompatibility C omplex class II b polymorphism influences gut microbiota composition and diversity. *Molecular ecology* 23, 4831–4845. doi: 10.1111/mec.12846

Bryant, M. (1970). Normal flora—rumen bacteria. *Am. J. Clin. Nutr.* 23, 1440–1450. doi: 10.1093/ajcn/23.11.1440

Burgos, S. A., Embertson, N. M., Zhao, Y., Mitloehner, F. M., Depeters, E. J., and Fadel, J. G. (2010). Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: relation of milk urea nitrogen to ammonia emissions. *J. Dairy Sci.* 93, 2377–2386. doi: 10.3168/jds.2009-2415

Burgos, S. A., Fadel, J. G., and Depeters, E. J. (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. doi: 10.3168/jds.2007-0299

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

Cantalapiedra-Hijar, G., Ortigues-Marty, I., Sepchat, B., Titgemeyer, E., and Bahloul, L. (2020). Methionine-balanced diets improve cattle performance in fattening young bulls fed high-forage diets through changes in nitrogen metabolism. *Br. J. Nutr.* 124, 273–285. doi: 10.1017/S0007114520001154

Carberry, C. A., Kenny, D. A., Han, S., Mccabe, M. S., and Waters, S. M. (2012). Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl. Environ. Microbiol.* 78, 4949–4958. doi: 10.1128/AEM.07759-11

Castillo, A., Kebreab, E., Beever, D., and France, J. (2000). A review of efficiency of nitrogen utilisation in lactating dairy cows and its relationship with environmental pollution. *J. Animal Feed Sci.* 9, 1–32. doi: 10.22358/jafs/68025/2000

Chebel, R. C., Santos, J. E., Reynolds, J. P., Cerri, R. L., Juchem, S. O., and Overton, M. (2004). Factors affecting conception rate after artificial insemination and pregnancy loss in lactating dairy cows. *Animal Reprod. Sci.* 84, 239–255. doi: 10.1016/j.anireprosci.2003.12.012

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

Cholewińska, P., Czyz, K., Nowakowski, P., and Wyrostek, A. (2020). The microbiome of the digestive system of ruminants: a review. *Animal Health Res. Rev.* 21, 3–14. doi: 10.1017/S1466252319000069

Chun, H., and Keleş, S. (2010). Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *J. R. Stat. Soc. Ser. B* 72, 3–25. doi: 10.1111/j.1467-9868.2009.00723.x

Correa-Luna, M., Donaghy, D., Kemp, P., Schutz, M., and López-Villalobos, N. (2021). Nitrogen use efficiency and excretion in grazing cows with high and low milk urea nitrogen breeding values. *Sustainability* 13, 9827. doi: 10.3390/su13179827

Coussens, P. M., Colvin, C. J., Rosa, G. J., Perez Laspiur, J., and Elftman, M. D. (2003). Evidence for a novel gene expression program in peripheral blood mononuclear cells from *Mycobacterium avium* subsp. paratuberculosis-infected cattle. *Infect. Immun.* 71, 6487–6498. doi: 10.1128/IAI.71.11.6487-6498.2003

De Mendiburu Delgado, F. (2009). *Una herramienta de análisis estadístico para la investigación* agrícola. Lima: Universidad Nacional de Ingeniería.

De Palma, G., Capilla, A., Nadal, I., Nova, E., Pozo, T., Varea, V., et al. (2010). Interplay between human leukocyte antigen genes and the microbial colonization process of the newborn intestine. *Curr. Issues Mol. Biol.* 12, 1–10. doi: 10.21775/cimb.012.001

Derakhshani, H., Plaizier, J. C., De Buck, J., Barkema, H. W., and Khafipour, E. (2018). Association of bovine major histocompatibility complex (BoLA) gene polymorphism with colostrum and milk microbiota of dairy cows during the first week of lactation. *Microbiome* 6, 1–18. doi: 10.1186/s40168-018-0586-1

Dias, J., Marcondes, M. I., Noronha, M. F., Resende, R. T., Machado, F. S., Mantovani, H. C., et al. (2017). Effect of pre-weaning diet on the ruminal archaeal, bacterial, and fungal communities of dairy calves. *Front. Microbiol.* 8, 1553. doi: 10.3389/fmicb.2017.01553

Ellis, S. A., and Ballingall, K. T. (1999). Cattle MHC: evolution in action? *Immunol. Rev.* 167, 159–168. doi: 10.1111/j.1600-065X.1999.tb01389.x

Emmanuel, B. (1980). Urea cycle enzymes in tissues (liver, rumen epithelium, heart, kidney, lung and spleen) of sheep (Ovis aries). *Comp. Biochem. Physiol. B Comp. Biochem.* 65, 693–697. doi: 10.1016/0305-0491(80)90181-9

Eschenlauer, S., Mckain, N., Walker, N., Mcewan, N., Newbold, C., and Wallace, R. (2002). Ammonia production by ruminal microorganisms and enumeration, isolation, and characterization of bacteria capable of growth on peptides and amino acids from the sheep rumen. *Appl. Environ. Microbiol.* 68, 4925–4931. doi: 10.1128/AEM.68.10.4925-4931.2002

Gardiner, C., Clough, T. J., Cameron, K. C., Di, H. J., Edwards, G., and De Klein, C. (2016). Potential for forage diet manipulation in New Zealand pasture ecosystems to mitigate ruminant urine derived N2O emissions: a review. *N. Zeal. J. Agric. Res.* 59, 301–317. doi: 10.1080/00288233.2016.1190386

Gleeson, M., Nieman, D. C., and Pedersen, B. K. (2004). Exercise, nutrition and immune function. *J. Sports Sci.* 22, 115–125. doi: 10.1080/0264041031000140590

Guliński, P., Salamończyk, E., and Młynek, K. (2016). Improving nitrogen use efficiency of dairy cows in relation to urea in milk: a review. *Anim. Sci. Pap. Rep.* 34, 5–24.

Hansen, P. J., Soto, P., and Natzke, R. P. (2004). Mastitis and fertility in cattle–possible involvement of inflammation or immune activation in embryonic mortality. *Am. J. Reprod. Immunol.* 51, 294–301. doi: 10.1111/j.1600-0897.2004.00160.x

Hartinger, T., Gresner, N., and Südekum, K.-H. (2018). Does intra-ruminal nitrogen recycling waste valuable resources? A review of major players and their manipulation. *J. Anim. Sci. Biotechnol.* 9, 1–21. doi: 10.1186/s40104-018-0249-x

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., and Janssen, P. H. (2015). Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Sci. Rep.* 5, 1–15. doi: 10.1038/srep14567

Hernandez-Sanabria, E., Guan, L. L., Goonewardene, L. A., Li, M., Mujibi, D. F., Stothard, P., et al. (2010). Correlation of particular bacterial PCR-denaturing gradient gel electrophoresis patterns with bovine ruminal fermentation parameters and feed efficiency traits. *Appl. Environ. Microbiol.* 76, 6338–6350. doi: 10.1128/AEM.01052-10

Honerlagen, H., Reyer, H., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., et al. (2021). Identification of genomic regions influencing N-metabolism and N-excretion in lactating holstein-friesians. *Front. Genet.* 12, 699550. doi: 10.3389/fgene.2021.699550

Hu, R., Zou, H., Wang, Z., Cao, B., Peng, Q., Jing, X., et al. (2019). Nutritional interventions improved rumen functions and promoted compensatory growth of growth-retarded yaks as revealed by integrated transcripts and microbiome analyses. *Front. Microbiol.* 10, 318. doi: 10.3389/fmicb.2019.00318

Hurley, D.J. (2014). "The relationship of immunity and reproduction in dairy cows," in *Proceedings of the 25th Annual Florida Ruminant Nutrition Symposium* (Gainesville, FL: University of Florida), 1.

Jahnel, R. E., Wittenburg, D., Mayer, M., Taeubert, H., and Reinsch, N. (2021). Estimation of genetic parameters of milk urea content. *Zuchtungskunde* 93, 157–168.

Jeong, J. Y., Ibrahim, M., Kim, M. J., So, K., Jeong, Y. D., Park, S., et al. (2017). Comparisons of extracellular matrix-related gene expression levels in different adipose tissues from Korean cattle. *Livestock Sci.* 198, 138–146. doi: 10.1016/j.livsci.2017.02.020

Jin, D., Zhao, S., Wang, P., Zheng, N., Bu, D., Beckers, Y., et al. (2016). Insights into abundant rumen ureolytic bacterial community using rumen simulation system. *Front. Microbiol.* 7, 1006. doi: 10.3389/fmicb.2016.01006

Jin, D., Zhao, S., Zheng, N., Beckers, Y., and Wang, J. (2018). Urea metabolism and regulation by rumen bacterial urease in ruminants: a review. *Ann. Anim. Sci.* 18, 303. doi: 10.1515/aoas-2017-0028

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Kolde, R., and Kolde, M. R. (2015). *Package 'pheatmap'*. Tartu: R Package 1, 790.

Kong, F., Gao, Y., Tang, M., Fu, T., Diao, Q., Bi, Y., et al. (2020). Effects of dietary rumen–protected Lys levels on rumen fermentation and bacterial community composition in Holstein heifers. *Appl. Microbiol. Biotechnol.* 104, 6623–6634. doi: 10.1007/s00253-020-10684-y

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

Kotz, A., Azevedo, P. A., Khafipour, E., and Plaizier, J. (2020). Effects of the dietary grain content on rumen and fecal microbiota of dairy cows. *Can. J. Anim. Sci.* 101, 274–286. doi: 10.1139/cjas-2020-0122

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/AEM.01043-13

Kvidera, S., Horst, E., Abuajamieh, M., Mayorga, E., Fernandez, M. S., and Baumgard, L. (2017). Glucose requirements of an activated immune system in lactating Holstein cows. *J. Dairy Sci.* 100, 2360–2374. doi: 10.3168/jds.2016-12001

Lapierre, H., Berthiaume, R., Raggio, G., Thivierge, M., Doepel, L., Pacheco, D., et al. (2005). The route of absorbed nitrogen into milk protein. *Anim. Sci.* 80, 11–22. doi: 10.1079/ASC41330011

Lavery, A., and Ferris, C. P. (2021). Proxy measures and novel strategies for estimating nitrogen utilisation efficiency in dairy cattle. *Animals* 11, 343. doi: 10.3390/ani11020343

Lewin, H. A., Russell, G. C., and Glass, E. J. (1999). Comparative organization and function of the major histocompatibility complex of domesticated cattle. *Immunol. Rev.* 167, 145–158. doi: 10.1111/j.1600-065X.1999.tb01388.x

Ling, J., and Armstead, I. (1995). The in vitro uptake and metabolism of peptides and amino acids by five species of rumen bacteria. *J. Appl. Bacteriol.* 78, 116–124. doi: 10.1111/j.1365-2672.1995.tb02831.x

Liu, W.-C., Guo, Y., An, L.-L., and Zhao, Z.-H. (2021). Protective effects of dietary betaine on intestinal barrier function and cecal microbial community in indigenous broiler chickens exposed to high temperature environment. *Environ. Sci. Pollut. Res.* 28, 10860–10871. doi: 10.1007/s11356-020-11326-6

Loubinoux, J., Valente, F. M., Pereira, I. A., Costa, A., Grimont, P. A., and Le Faou, A. E. (2002). Reclassification of the only species of the genus Desulfomonas, Desulfomonas pigra, as Desulfovibrio piger comb. *nov. Int. J. Syst. Evolut. Microbiol.* 52, 1305–1308.: doi: 10.1099/00207713-52-4-1305

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. doi: 10.1186/s13059-014-0550-8

Mcmurdie, P. J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* 10, e1003531. doi: 10.1371/journal.pcbi.1003531

Meehan, C. J., and Beiko, R. G. (2014). A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. *Genome Biol. Evolut.* 6, 703–713. doi: 10.1093/gbe/evu050

Miglior, F., Sewalem, A., Jamrozik, J., Bohmanova, J., Lefebvre, D. M., and Moore, R. K. (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. doi: 10.3168/jds.2006-487

Moore, D. A., Overton, M. W., Chebel, R. C., Truscott, M. L., and Bondurant, R. H. (2005). Evaluation of factors that affect embryonic loss in dairy cattle. *J. Am. Vet. Med. Assoc.* 226, 1112–1118. doi: 10.2460/javma.2005.226.1112

Müller, C. B. M., Görs, S., Derno, M., Tuchscherer, A., Wimmers, K., Zeyner, A., et al. (2021). Differences between Holstein dairy cows in renal clearance rate of urea affect milk urea concentration and the relationship between milk urea and urinary nitrogen excretion. *Sci. Total Environ.* 755, 143–198. doi: 10.1016/j.scitotenv.2020.143198

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

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., et al. (2013). Package 'vegan'. *Commun. Ecol.* 2, 1–295. Available online at: http://cran.r-project.org/

Pacífico, C., Petri, R. M., Ricci, S., Mickdam, E., Wetzels, S. U., Neubauer, V., et al. (2021). Unveiling the bovine epimural microbiota composition and putative function. *Microorganisms* 9, 342. doi: 10.3390/microorganisms9020342

Paradis, F., Yue, S., Grant, J., Stothard, P., Basarab, J., and Fitzsimmons, C. (2015). Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. *J. Animal Sci.* 93, 3331–3341. doi: 10.2527/jas.2015-8975

Pasquariello, R. (2015). *Molecular Characterization of Factors Which Can Influence the Reproductive Success of Cattle*. Milano : Universitá degli Studi di Milano, Dipartimento di Scienze Agrarie e Ambientali.

Patra, A. K., and Aschenbach, J. R. (2018). Ureases in the gastrointestinal tracts of ruminant and monogastric animals and their implication in urea-N/ammonia metabolism: a review. *J. Adv. Res.* 13, 39–50. doi: 10.1016/j.jare.2018.02.005

Penner, G., Steele, M., Aschenbach, J., and Mcbride, B. (2011). Ruminant nutrition symposium: molecular adaptation of ruminal epithelia to highly fermentable diets. *J. Anim. Sci.* 89, 1108–1119. doi: 10.2527/jas.2010-3378

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protocols* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Peterson, D. A., Mcnulty, N. P., Guruge, J. L., and Gordon, J. I. (2007). IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell Host Microbe* 2, 328–339. doi: 10.1016/j.chom.2007.09.013

Plaizier, J. C., Li, S., Tun, H. M., and Khafipour, E. (2017). Nutritional models of experimentally-induced subacute ruminal acidosis (SARA) differ in their impact on rumen and hindgut bacterial communities in dairy cows. *Front. Microbiol.* 7, 2128. doi: 10.3389/fmicb.2016.02128

Poehlein, A., Schneider, D., Soh, M., Daniel, R., and Seedorf, H. (2018). Comparative genomic analysis of members of the genera Methanosphaera and Methanobrevibacter reveals distinct clades with specific potential metabolic functions. *Archaea* 2018. doi: 10.1155/2018/7609847

Razmi, N., Ali Jelodar, G., Nazifi, S., and Dehghani, A. (2005). Arginase status in cattle reproductive system. *Veterinarski arhiv* 75, 31–38.

Rohart, F., Gautier, B., Singh, A., and Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* 13, e1005752. doi: 10.1371/journal.pcbi.1005752

Rosendahl, J. (2015). *Untersuchung über die beteiligung von nicht-selektiven kationenkanälen am ammoniaktransport über das pansenepithel des rindes* (Dissertation). Berlin: Freie Universität Berlin, Institut für Veterinär-Physiologie des Fachbereichs Veterinärmedizin.

Rupp, R., and Boichard, D. (2003). Genetics of resistance to mastitis in dairy cattle. *Vet. Res.* 34, 671–688. doi: 10.1051/vetres:2003020

Russell, J. B., Muck, R. E., and Weimer, P. J. (2009). Quantitative analysis of cellulose degradation and growth of cellulolytic bacteria in the rumen. *FEMS Microbiol. Ecol.* 67, 183–197. doi: 10.1111/j.1574-6941.2008.00633.x

Sales, M., Lucas, F., and Blanchart, G. (2000). Effects of ammonia and amino acids on the growth and proteolytic activity of three species of rumen bacteria: prevotella albensis, Butyrivibrio fibrisolvens, and Streptococcus bovis. *Curr. Microbiol.* 40, 380–386. doi: 10.1007/s002840010074

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Spek, J. W., Dijkstra, J., Van Duinkerken, G., and Bannink, A. (2012). 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. doi: 10.1017/S0021859612000561

Steele, M. A., Penner, G. B., and Chaucheyras-Durand, F. (2016). Development and physiology of the rumen and the lower gut: targets for improving gut health. *J. Dairy Sci.* 99, 4955–4966. doi: 10.3168/jds.2015-10351

Suen, G., Stevenson, D. M., Bruce, D. C., Chertkov, O., Copeland, A., Cheng, J.-F., et al. (2011). Complete genome of the cellulolytic ruminal bacterium *Ruminococcus albus* . *Am. Soc. Microbiol.* 193, 5574–5575. doi: 10.1128/JB.05621-11

Sun, H. Z., Zhou, M., Wang, O., Chen, Y., Liu, J. X., and Guan, L. L. (2020). Multi-omics reveals functional genomic and metabolic mechanisms of milk production and quality in dairy cows. *Bioinformatics* 36, 2530–2537. doi: 10.1093/bioinformatics/btz951

Tan, P., Liu, H., Zhao, J., Gu, X., Wei, X., Zhang, X., et al. (2021). Amino acids metabolism by rumen microorganisms: nutrition and ecology strategies to reduce nitrogen emissions from the inside to the outside. *Sci. Total Environ.* 800, 149–596. doi: 10.1016/j.scitotenv.2021.149596

Tannock, G. W., and Liu, Y. (2020). Guided dietary fibre intake as a means of directing short-chain fatty acid production by the gut microbiota. *J. R. Soc. N. Zeal.* 50, 434–455. doi: 10.1080/03036758.2019.1657471

Te Pas, M. F., Veldkamp, T., De Haas, Y., Bannink, A., and Ellen, E. D. (2021). Adaptation of livestock to new diets using feed components without competition with human edible protein sources—a review of the possibilities and recommendations. *Animals* 11, 2293. doi: 10.3390/ani11082293

Toivanen, P., Vaahtovuo, J., and Eerola, E. (2001). Influence of major histocompatibility complex on bacterial composition of fecal flora. *Infect. Immun.* 69, 2372–2377. doi: 10.1128/IAI.69.4.2372-2377.2001

Trabi, E. B., Seddik, H.-E., Xie, F., Wang, X., Liu, J., and Mao, S. (2020). Effect of pelleted high-grain total mixed ration on rumen morphology, epithelium-associated microbiota and gene expression of proinflammatory cytokines and tight junction proteins in Hu sheep. *Anim. Feed Sci. Technol.* 263, 114453. doi: 10.1016/j.anifeedsci.2020.114453

Tun, H. M., Li, S., Yoon, I., Meale, S. J., Azevedo, P. A., Khafipour, E., et al. (2020). Saccharomyces cerevisiae fermentation products (SCFP) stabilize the ruminal microbiota of lactating dairy cows during periods of a depressed rumen pH. *BMC Vet. Res.* 16, 237. doi: 10.1186/s12917-020-02437-w

Uwizeye, A., De Boer, I. J. M., Opio, C. I., Schulte, R. P. O., Falcucci, A., Tempio, G., et al. (2020). Nitrogen emissions along global livestock supply chains. *Nat. Food* 1, 437–446. doi: 10.1038/s43016-020-0113-y

Vereinigte Informationssysteme Tierhaltung (VIT). (2020). *Beschreibung der Zuchtwertschätzung für alle Schätzmerkmale bei den Milchrinderrassen für die vit mit der Zuchtwertschätzung beauftragt ist*. Available online at: https://www.vit.de/fileadmin/DE/Zuchtwertschaetzung/Zws_Bes_deu.pdf. (accessed January 01, 2020).

Wallace, R., and Brammall, M. L. (1985). The role of different species of bacteria in the hydrolysis of protein in the rumen. *Microbiology* 131, 821–832. doi: 10.1099/00221287-131-4-821

Wang, L., Liu, K., Wang, Z., Bai, X., Peng, Q., and Jin, L. (2019). Bacterial community diversity associated with different utilization efficiencies of nitrogen in the gastrointestinal tract of goats. *Front. Microbiol.* 10, 239. doi: 10.3389/fmicb.2019.00239

Wang, M.-Z., Shen, J.-Z., Liu, Y., Dong, S.-H., Yu, L.-H., and Wang, H.-R. (2011). Effects of dietary protein sources on bacterial and protozoal community structure in goat rumen. *Chin. J. Anim. Sci.* 3:1–12. doi: 10.1371/journal.pone.0221815

Wang, Y., Nan, X., Zhao, Y., Jiang, L., Wang, M., Wang, H., et al. (2021). Rumen microbiome structure and metabolites activity in dairy cows with clinical and subclinical mastitis. *J. Anim. Sci. Biotechnol.* 12, 1–21. doi: 10.1186/s40104-020-00543-1

Weiss, G., Goldsmith, L. T., Taylor, R. N., Bellet, D., and Taylor, H. S. (2009). Inflammation in reproductive disorders. *Reprod. Sci.* 16, 216–229. doi: 10.1177/1933719108330087

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

Zhang, N., Teng, Z., Li, P., Fu, T., Lian, H., Wang, L., et al. (2021). Oscillating dietary crude protein concentrations increase N retention of calves by affecting urea-N recycling and nitrogen metabolism of rumen bacteria and epithelium. *PLoS ONE* 16, e0257417. doi: 10.1371/journal.pone.0257417

Zhang, T., Mu, Y., Zhang, D., Lin, X., Wang, Z., Hou, Q., et al. (2019). Determination of microbiological characteristics in the digestive tract of different ruminant species. *Microbiologyopen* 8, e00769. doi: 10.1002/mbo3.769

11.4 Microbial signatures inferred from association with genomic breeding values for milk urea concentration (MU) and their relationship to proxies for N utilisation efficiency (NUE) in Holstein cows (4th study)

Running head: Rumen microbiota of cows selected for Milk urea

Microbial signatures inferred from association with genomic breeding values for milk urea concentration (MU) and their relationship to proxies for N-utilisation efficiency (NUE) in Holstein cows

Hanne Honerlagen[1], Henry Reyer[1], Ibrahim Abou-Soliman[1], Dierck Segelke[2], Siriluck Ponsuksili[1], Nares Trakooljul[1], Norbert Reinsch[4], Björn Kuhla[3], Klaus Wimmers[1, 5]*

[1]Research Institute for Farm Animal Biology, Institute of Genome Biology, Dummerstorf, Germany
[2]IT-Solutions for Animal Production, Vereinigte Informationssysteme Tierhaltung w.V. (vit), 27283 Verden, Germany
[3]Research Institute for Farm Animal Biology, Institute of Nutritional Physiology "Oskar Kellner", Dummerstorf, Germany
[4]Research Institute for Farm Animal Biology, Institute of Genetics and Biometry, Dummerstorf, Germany
[5]University of Rostock, Faculty of Agricultural and Environmental Sciences, Rostock, Germany
[*] **Correspondence:** Klaus Wimmers; wimmers@fbn-dummerstorf.de; phone: +49 38208 68-600

Abstract:

Increasing the nitrogen (N)-utilisation efficiency (NUE) of dairy cows by breeding selection would offer advantages from a nutritional, environmental and economic perspective. Since data collection of NUE phenotypes is not feasible in large cow cohorts, the cow individual milk urea concentration (MU) has been suggested as indicator trait. In regard of the symbiotic interplay between dairy cows and their rumen microbiome, the individual MU was thought to be influenced by host genetics and by the rumen microbiome, the latter in turn being partly attributed to host genetics. In order to enhance the knowledge of MU as indicator trait for NUE, this study aimed to identify differential abundant rumen microbial genera between Holsteins with divergent genomic breeding values for MU (GBVMU; GBV_{HMU} vs. GBV_{LMU}). The microbial genera identified were further investigated for their correlations to MU and seven additional NUE-associated traits in urine, milk and faeces in 358 lactating Holsteins. Statistical analysis of microbial 16S rRNA amplicon sequencing data revealed significantly higher abundances of the ureolytic genus *Succinivibrionaceae UCG-002* in GBV_{LMU} cows, whereas GBV_{HMU} animals hosted higher abundances of *Clostridia unclassified* and *Desulfovibrio*. The entire discriminating ruminal signature of 24 microbial taxa included a further three genera of the *Lachnospiraceae* family that revealed significant correlations to MU values and were therefore proposed as considerable players in the GBVMU – microbiome – MU axis. The significant correlations of *Prevotellaceae UCG-003, Anaerovibrio, Blautia* and *Butyrivibrio* abundances with MU measurements, milk nitrogen (MN) and the N content in faeces (FaecN) suggested their contribution to genetically determined N-utilisation in Holstein cows. The microbial genera identified might be considered for future breeding programs in order to enhance NUE in dairy herds.

Keywords: Rumen microbiota[1], genomic breeding value milk urea[2], nitrogen utilisation efficiency[3], host – rumen microbiota – trait[4], dairy cow[5]

Introduction

The symbiotic relationship between the dairy cow and its rumen microbiome enables the utilisation of non-protein nitrogen (i.e. urea and ammonia) and increases the biological value of the dietary protein by rumen microbial protein synthesis (Bryant, 1970, Huber and Kung Jr, 1981, Tan et al., 2021). Although ruminal N-conversion in general increases the N-availability for the host, the N-utilisation efficiency (NUE) of dairy cows – defined as the ratio of gram N in milk to gram N intake –is on average only 25% (Calsamiglia et al., 2010). Non-utilised dietary-N and N of endogenous origin is excreted via milk, urine and faeces (Abdoun et al., 2006, Reed et al., 2015, Tan et al., 2021). Whereas faecal N accounts quantitatively for the largest part of N-excretion and is therefore most considerable for

nutrient losses, the urinary N-excretion – with urea as the major N-excretion metabolite – contributes remarkably to N-emissions from dairy farms (Dijkstra et al., 2011, Spanghero and Kowalski, 2021). A more efficient N-utilisation by the symbiotic interplay between dairy cows and their rumen microbiome would reduce the excretion of non-utilised N and provide a higher nutrient availability, accompanied by economic and ecological advantages (Spanghero and Kowalski, 2021, Tan et al., 2021). Interestingly, NUE was found to vary widely between individual cows, with ranges from 16% to 36 % across diets, N-intake levels, breeds, lactation numbers and lactation stages (Calsamiglia et al., 2010, Powell et al., 2010). Recently, early lactating dairy cows were even phenotyped with NUEs ranging from 9.7% to 81.7% (Grelet et al., 2020, Bergen, 2021, Jahnel et al., 2021, Te Pas et al., 2021).

To overcome this problem, various studies proposed the individual milk urea concentration (MU) as an indicator trait (Nousiainen et al., 2004, Guliński et al., 2016, Bobbo et al., 2020, Lavery and Ferris, 2021). MU is known to be moderately correlated with the urinary urea concentration (UU) and was therefore thought to depict N-emissions, deriving from UU (Gonda and Lindberg, 1994, Burgos et al., 2010, Beatson et al., 2019). Moreover, MU has been attributed a predictive power about cow individual NUE (Nousiainen et al., 2004, Huhtanen et al., 2015, Munyaneza et al., 2017, Bobbo et al., 2020, Lavery and Ferris, 2021). Due to its routinely obtainment by monthly milk records, huge data sets of individual MU values are available. The moderate heritability of MU as well as the weak or absent genetic correlations to milk performance parameters further support MU' s potential as an indicator trait for breeding programs (Wood et al., 2003, Miglior et al., 2007). However, the success of breeding selection by using an indicator trait strongly depends on its prediction accuracy for the real trait (MU – NUE). Thus, increasing the knowledge about the relationship between MU genetics and NUE phenotypes is a necessary step before MU can be seriously considered for breeding strategies.

The rumen microbial community is known as the major determinant of the N metabolism in ruminants (Bach et al., 2005). Although the rumen microbiome is strongly influenced by the diet (Loor et al., 2016), recent studies identified significant host genetic effects on rumen microbial abundances (Difford et al., 2018, Pérez-Enciso et al., 2021, Saborío-Montero et al., 2021). Genomic breeding selection for MU is therefore hypothesized to affect the host-trait axis and the triangular host – rumen microbiome – trait axis, which collectively influence individual NUE phenotypes.

In order to better understand the link between MU genetics and individual proxies of NUE, the current study aimed to explore both axes by (i) the identification of differential microbial genera abundances between dairy cows with divergent genomic breeding values for MU (GBV_{LMU} vs. GBV_{HMU}) and (ii) by investigating these rumen microbial genera identified in (i) as potential microbial signatures of GBVMU selection for their relationship to proxies of individual NUE.

Material and Methods

The following description of data collection, statistical analyses and the analysis strategy is visualised in Figure 1.

Cow population and sampling

The cow population and the experimental trial has been described previously (Honerlagen et al., 2021). In brief, 371 lactating Holstein-Friesian cows were sampled for milk, urine, faeces and rumen fluid in a practice-operating dairy farm. All cows were fed a total mixed ration which comprised 7.26 MJ NEL/kg dry matter (DM), 15.4% crude protein (CP) and a ruminal nitrogen balance of -0.9 g N/kg DM. Milking and feeding was carried out twice daily. Urine, faeces and rumen fluids were obtained once from each cow in a maximum time distance of 24h to milk samples, which were obtained by the monthly milk record procedure as pooled sample from morning and evening milking. Monthly milk recording data for one entire lactation (14 months) including protein, fat, lactose and milk yield was available for all individuals of the herd. In order to sample urine, faeces and rumen fluids, cows were fixed in a feeding fence directly after the morning or evening milking. Urine was collected after stimulating massages or spontaneous miction, immediately cooled on ice and stored at -20 °C. Faeces samples (approx. 200 mg per cow) were withdrawn by rectal removal and stored at -20°C. Rumen fluids were obtained by oral stomach tubing. Therefore, each cow was fixed with a rope in the feeding fence, a stomach tube (Hauptner-Herberholz GmbH & Co. KG, Solingen, Germany) was warmed up in water and intubated into the rumen. After discarding saliva, rumen fluids were transferred into 50-ml falcon tubes and stored at ambient temperature (1-7°C) for 2 h maximum before long-time storage at -20°C. The stomach tube was intensively rinsed with water before the next cow was sampled.

Genomic breeding value estimation for MU (GBVMU)

For all cows a genomic breeding value for MU (GBVMU) was estimated by vit Verden (Vereinigte Informationssysteme Tierhaltung, Verden, Germany) by applying the established estimation model for somatic cell score (https://www.vit.de/fileadmin/DE/Zuchtwertschaetzung/Zws_Bes_deu.pdf, access 20.01.2020). Only cows with GBVMUs and all phenotypes were included in the statistical analyses, resulting in a data set of n=358 lactating Holsteins.

Based on GBVMUs the population was grouped into the following three categories: GBV_{HMU} (GBVMU ≤ 85, n=30, MU phenotype= 226.97mg/l ± 19.88 (mean MU ± SD of the group)), GBV_{MED} (GBVMU = 86 – 115, n=299; MU phenotype= 191.95 mg/l ± 21.24) and GBV_{LMU} (GBVMU ≥ 115, n=29, MU phenotype= 161.54 mg/l ± 19.10). The extreme cow groups (GBV_{HMU}, GBV_{LMU}) represented the top and bottom 15% for GBVMU of the population separated by three standard deviations between GBVMU values. The cow groups assigned to GBVMU were similar in parity (3, 3, 3; mean GBV_{HMU}, $GBV_{LMU,}$ GBV_{MED}), lactation stages (198 days in milk (DIM), 203 DIM, 199 DIM) and body weights (681 kg, 686 kg, 684 kg).

Sample analysis and settings of NUE proxies

The analyses of milk and urine samples were previously described (Honerlagen et al., 2021). In brief, the State Control Federation of Mecklenburg-Western Pomerania provided the analyses of milk protein, milk fat, lactose and milk urea, the latter being analysed by mid infrared spectroscopy (CombiFoss 7, Foss, Hilleroed, Denmark). Total milk N was determined by MQD (Qualitätsprüfungs- und Dienstleistungsgesellschaft Mecklenburg-Vorpommern, Güstrow, Germany) with the Kjeldahl method and converted into milk CP (g/100g milk), using the factor 6.38. The urine samples were analysed photometrically for urea concentration at a ABX Pentra C400 clinical chemistry analyser (HORIBA Europe GmbH, Oberursel, Germany). Faeces samples were dried, grinded and subsequently analysed for the total N-content, calculated as percentage of DM, using a “vario MAX” element analyzer (Elementar; Langenselbold, Germany). The rumen fluids were analysed for the microbial composition by 16S rRNA amplicon sequencing. Microbial DNA extraction, 16S rRNA amplicon sequencing and data preparation followed the description in (Honerlagen et al., 2022). Briefly, microbial DNA was extracted utilising the PowerLyzer PowerSoil DNA isolation kit (QIAGEN, Hilden,

Germany), the V4 region of the 16S rRNA gene was targeted via PCR and the PCR products were subsequently sequenced on HiSeq2500 (Illumina, San Diego, CA) with 250 bp paired-end reads. After excluding nine samples due to low read depth the remaining data set was subsampled to 120 521 reads per sample (McMurdie and Holmes, 2014).

Proxies of NUE generated from milk, urine and faeces samples comprised repeatedly measured and one-time measured traits. Repeatedly measured traits were generated by averaging individual milk record data of one lactation (14 months lactation duration). One-time measured traits were recorded as part of the collection procedure of rumen, urine and faecal samples. Repeatedly measured proxies of NUE utilised in this study are: Milk urea concentration (MU_{lac}), milk yield (MY_{lac}), milk urea yield (MUY_{lac}= MU_{lac}* MY_{lac} and milk protein yield (MPY_{lac} = milk protein percentage (MP_{lac})*MY_{lac}). One-time measured traits are: MU, urinary urea concentration (UU), milk protein percentage (MP), milk CP as parameter for total milk nitrogen concentration (MN) and faecal N-concentration (FaecN).

Statistical analyses

Milk, urine and faeces phenotypes were tested for significant differences between the extreme GBVMU groups (GBV_{HMU} vs. GBV_{LMU}) by utilising Student's t-test and further investigated for phenotypic correlations among traits in the whole cow population (n=358).

The microbial data was initially analysed at OTU level in the whole population (n=358) by multidimensional scaling (NMDS) based on a Bray-Curtis dissimilarity matrix using 'vegan' R package (Oksanen et al., 2013). Moreover, PERMANOVA analysis was performed based on the dissimilarity matrix to test differences between GBV_{HMU}, GBV_{MED} and GBV_{LMU} groups. Inverse Simpson indices were calculated to assess alpha diversity with 'agricolae' R package (De Mendiburu Delgado, 2009).

Further analyses were conducted at genus level considering genera with more than 30 counts in at least a third of the entire cow population (remaining data set: 123 genera). Differences in taxa abundance between the extreme GBVMU groups were analysed using Wald test implemented in the 'DESeq2' R package (Love et al., 2014). The milk fat content was included as fixed effect in the statistical model accounting for differences in feed intake between the cows. Genera were defined as significantly differentially abundant (DAG) at *p*-values < 0.05. Furthermore, a sparse Partial Least Squares Discriminant Analysis (sPLS-DA) was applied on the filtered and variance-stabilised transformed count data set (123 genera) utilising the R package 'mixOmics' (version 6.6.2, available at: http://mixomics.org) (Rohart et al., 2017). SPLS-DA was utilised to identify the most discriminative microbial genera that mainly distinguish the rumen profiles of extreme GBVMU groups. The analysis considered two components of microbial features, with the 10 most distinctive genera in each of the components.

The microbial genera identified by 'DESeq2' (DAG) and 'sPLS-DA' were considered as microbial signatures of GBVMU and further investigated for their correlation to the NUE-associated traits (milk, urine, faeces) in the whole cow population (n= 358). Therefore, a Pearson correlation analysis between the sample-specific genus abundance and each trait was conducted and tested for significance by utilising 'stats' package in R. Significance was determined at adjusted *p*-values < 0.05 (Benjamini-Hochberg). Most prominent microbe-trait correlations considering correlation coefficients, significance and literature research were identified and visualised with 'ggplot' in R.

Results

Phenotypic characterisation of cows assigned to GBVMU

The NUE-associated traits are documented in Table 1 showing means and standard deviations for the entire herd, as well as for GBV_{HMU} and GBV_{LMU} cow groups.
GBV_{LMU} cows showed significantly lower MU values than GBV_{HMU} cows, which was evident at sampling time point (MU) and across lactation (MU_{lac}). GBV_{LMU} cows also excreted significantly less absolute MU per lactation, indicated by MUY_{lac}, even though they yielded significantly more milk (MY_{lac}) than

GBV_{HMU} cows. Furthermore, GBV_{LMU} had significantly lower UU than GBV_{HMU} cows, whereas MN, MP, MPY_{lac} and FaecN did not significantly differ between groups.
Correlations among the NUE-associated traits in the whole cow population are reported in Table 2. MU_{lac} and MU were positively correlated (r=0.63, p<0.05). UU revealed moderate positive correlations to MU_{lac} (r=0.25) and MU (r=0.26) and a weak positive correlation to MN (r=0.14). FaecN did not show any significant correlations to milk and urine traits. Although a moderate correlation between the repeatedly measured MPY_{lac} and MU_{lac} was observed (r=0.25), there was no relation between MPY_{lac} and one-time measured MU. However, one-time measured MU correlated positively with the one-time measured MP (r=0.27).

Results microbiome

The filtered and subsampled microbial data set in the entire cow population comprised 1064 genera, which were taxonomically assigned to 426 families. Analysis at OTU level revealed neither a distinct clustering of the overall microbial communities (Figure 2A) nor differences in the alpha diversity between GBVMU groups (GBV_{HMU}, GBV_{MED}, GBV_{LMU}; Figure 2B).

Analyses at genus level uncovered 13 DAG between GBV_{HMU} and GBV_{LMU} cows in the filtered data set (Table 3). The genera with the highest abundance among the DAG were *Lachnospiraceae NK3A20 group* and *Muribaculaceae ge*, both with a higher occurrence in GBV_{HMU} cows compared to GBV_{LMU} animals. Two additional genera assigned to the *Lachnospiraceae* were identified as DAG. Furthermore, *Veillonellaceae unclassified*, *Succinivibrionaceae UCG-002* and *Blautia* abundances indicated substantial differences between the divergent GBVMU groups. Whereas *Veillonellaceae unclassified* was observed 2.5 times more in GBV_{HMU} compared to GBV_{LMU}, *Succinivibrionaceae UCG-002* and *Blautia* were significantly more abundant in GBV_{LMU} with fold changes of 2.7 and 2.1, respectively. The significantly higher abundance of *Desulfovibrio* in GBV_{HMU} was supported by the lowest *p*-value in the data set, whereas *Clostridia unclassified* might constitute a further prominent DAG with higher prevalence in the rumen of GBV_{HMU} cows.

Although sPLS-DA did not achieve a complete distinction of GBV_{HMU} and GBV_{LMU} cows' microbial profiles, the abundances of the selected microbial features partly separated the rumen profiles of the divergent GBVMU groups (Figure 3A). *Lachnospiraceae NK3A20 group*, which was identified as most abundant DAG in the 'DeSeq2' analysis, was also selected as most important driver of component 1, accounting for higher abundances in GBV_{HMU} (Figure 3B, Table 3). Eight further DAG were also identified by sPLS-DA, whereas *WCHB1-41 ge* and *Bacteroidales BS11 gut group ge* were uncovered exclusively by sPLS-DA and represented the two most important features of component 2 (Figure 3C). Both genera occupied higher abundances in GBV_{HMU} cows and covered mean relative abundances of 0.2378% and 0.1445% in the whole cow population.

The correlation analysis between microbial abundances and NUE-associated traits (milk, urine, faeces) in the entire cow population (n=358) was conducted with the 24 microbial genera related to GBVMU obtained from the joint lists of 'DESeq2' and 'sPLS-DA' (Table 4; Suppl.1). Most prominent genus abundance – trait correlations are visualised in Figure 4 (A-F).

The abundances of two genera of *Lachnospiraceae* family, namely the "*XPB1014 group*" and "*AC2044 group*", as well as *Prevotellaceae NK3B31 group* and *Butyrivibrio* revealed considerable negative correlation coefficients to MU_{lac} (r^2 = -0.16 to -0.13). *Butyrivibrio* and *Lachnospiraceae XPB1014 group* abundances showed further notable negative correlations to MUY_{lac} (r^2 = -0.13 and r^2 = -0.14, respectively), but positive relations to FaecN (r^2 = 0.24; r^2 = 0.21). Another six genera were significantly correlated with MU (r^2 = > abs (0.15); adj. p < 0.05). Whereas *Erysipelotrichaceae UCG-007* and *Clostridia unclassified* abundances correlated negatively with MU, higher abundances of *Prevotellaceae UCG-003*, *Blautia* and *Anaerovibrio* were observed alongside higher MU values. The latter three genera were further found to be positively correlated to MN and FaecN. *Blautia* abundances revealed thereby the most remarkable relationships (MN: r^2 = 0.2; FaecN: r^2 = 0.29), but was only barely observed (0.0344% mean relative abundance). Moreover, the abundances of three

genera (*WCHB1-41 ge*, *CAG-352* and *Bacteroidales BS11 gut group ge*), which contributed to the discrimination of extreme GBVMU groups in sPLS-DA, were accompanied by higher FaecN excretion and lower MUY_{lac} and MY_{lac} values. Furthermore, the abundances of *Desulfovibrio* and *Veillonellaceae unclassified* were negatively correlated with MN, MP and FaecN. Although both genera were identified as DAG, their abundances were not correlated to MU_{lac} or MU phenotypes considering the entire cow herd. *Anaerolineae unclassified* abundances revealed the only considerable correlation with UU values.

Discussion:

This study was conducted under the hypothesis that due to the symbiotic interplay between dairy cows and their rumen microbiome, individual MU is affected by host-genetics and the rumen microbiome, which in term is partly attributed to host genetics. Thus, differences in GBVMU should be displayed by different abundances of rumen microbes playing a role for the individual N-utilisation and N-excretion. Specific microbial genera were identified as potential rumen microbial signatures related to GBVMU. These genera were further investigated for their relationship to proxies of NUE in milk, urine and faeces in order to assess their potential influence on cow individual NUE phenotypes.

All NUE associated phenotypes ranged in standard norms, assuring suitable dietary-N and energy supply in a high producing herd (Gonda and Lindberg, 1994, Sørensen et al., 2003, Ruska and Jonkus, 2014). The phenotypic characterisation of milk parameters revealed that GBV_{LMU} cows excrete significantly less urea via milk (MU_{lac}, MU and MUY_{lac}), but yielded more milk (MY_{lac}) than GBV_{HMU}. However, in accordance with various studies that confirmed the absence of genetic and phenotypic correlations between MU and milk performance parameters in large cow populations (Wood et al., 2003, Miglior et al., 2007), no phenotypic correlations between MU_{lac} and MU with MY_{lac} were identified by the present study. Thus, higher MY_{lac} in GBV_{LMU} cows do not indicate that GBVMU selection would concurrently enhance MY_{lac}. However, our results strengthen the hypothesis that GBVMU selection would at least not negatively influence MY_{lac}. Moreover, even though GBV_{LMU} cows were phenotyped with higher MY_{lac} than GBV_{HMU}, they excreted significantly less absolute MU per lactation (MUY_{lac}= MU_{lac}*MY_{lac}). These findings indicate the high potential to substantially reduce MU_{lac} by GBVMU breeding selection.
Urea – as the quantitatively most occurring N-metabolite in urine – is also known as the most variable urinary N-fraction (Bristow et al., 1992), which might explain the high standard deviations in our study. Nonetheless, lower UU in GBV_{LMU} compared to GBV_{HMU} was significantly evident. The moderate correlations between UU and MU (MU_{lac} and MU) are concordant with various studies and strengthen the general assumption that selection on GBV_{LMU} cows would reduce UU and thus N-emissions (Burgos et al., 2010, Guliński et al., 2016, Bergen, 2021, Te Pas et al., 2021).
Although GBV_{LMU} cows had significantly lower urea concentrations in milk and urine they did not occupy significantly different FaecN phenotypes compared to GBV_{HMU} cows. This observation is in accordance with (Arunvipas et al., 2008) who assured absent associations between MU and faecal N concentration by multiple measurements in 79 lactating dairy herds. These findings suggest the independency of GBVMU selection to the excretion of undigested N in faeces. However, it should be considered that data collection of absolute urinary and faecal N-excretion was not possible in the present study, due to the free-stall housing of the herd, the large sample size and the sample obtainment within the daily farming routine. Hence, some NUE associated phenotypes were depicted relative to the normalized sampling volume (i.e. UU and FaecN). The determination of absolute N-losses by exact calculations of N-input, N-deposit (i.e. milk and muscle protein) and N-excretion would enhance the knowledge of phenotypes deriving from GBVMU selection.

N compounds in milk and urine origin from the cows' N-pool as a result of N- absorption in the rumen and intestine, whereas faecal N is mainly determined by non- absorbed N. Consequently, a potential influence of specific rumen microbial genera on proxies of NUE demands for differentiation between milk and urine traits compared to faecal N excretion.

In general, a high ruminal fermentation rate of feed-N compounds yielding high amounts of absorbable N-metabolites (i.e. peptides and AA) at the small intestine enhances the N-usability for the cow,

decreases the ruminal and the cow's (blood) non-protein-N (NPN) pool and thus reduces N-losses via milk and urine. A specific rumen microbial effect on milk and urinary N-excretion would therefore be conceivable, if a genus (i) influences the ruminal amount of NH_3 or NH^{4+} by its deamination activity (impact on ruminal NPN-pool), (ii) determines the diffusion and utilisation of blood urea in the rumen by its ureolytic activity (impact on cow's NPN pool), or (iii) influences the transport of NH_3, NH^{4+} and urea across the rumen epithelium (impact on ruminal and cow's NPN-pool).

Faecal N-excretion originates from undigested dietary N, non-absorbed microbial protein, and to a small amount from endogenous N (Stallcup et al., 1975). The faecal N-content is therefore mainly determined by the total amount of N absorbed by the cow, independent of the N-absorption form (NH_3 or NH^{4+} via rumen epithelia; AA and peptides via small intestine membrane). Microbial influence on FaecN phenotypes might therefore be more due to (i) the initial breakdown of dietary protein-N by proteolytic activity in the rumen, (ii) the ruminal N-fermentation activity, which determines the amount of N incorporation into microbial protein or (iii) an effect on the dietary passage rate, which impacts the cow's intestinal barrier absorption capacity of peptides and AA.

The results of the overall microbial composition indicated that GBVMU selection is neither accompanied by specific ruminotypes nor would the microbial diversity been substantially increased or decreased. These findings are in accordance with low to moderate heritability estimations of microbial features in dairy cattle (Difford et al., 2018) and might be explained by the fact that the rumen microbial composition is mainly influenced by environmental factors (i.e. feed components) and only to a certain extent by host genetics (Yáñez-Ruiz et al., 2015). Nonetheless, fourteen DAG were identified between the extreme GBVMU groups. Interestingly, seven thereof were also identified to distinguish the rumen profiles of another cow population grouped for high and low MU predisposition in our previous study (Honerlagen et al., 2022). Specifically, a higher abundance of the ureolytic genus *Succinivibrionaceae UCG-002* was identified in GBV_{LMU} cows, whereas GBV_{HMU} cows hosted significantly more abundances of *Clostridia unclassified* and *Desulfovibrio*, which became prominent as hyper-ammonia producing bacteria (HAB) species (Paster et al., 1993, Bento et al., 2015, Hartinger et al., 2018, Honerlagen et al., 2021, Libera et al., 2021). Ureolytic bacteria are known to enhance the diffusion of blood urea into the rumen, which reduces the cow's blood NPN-pool, facilitates the ruminal N-incorporation into microbial protein and thus reduces N-losses (Rosendahl, 2015). Accordingly, ureolytic genera were identified to facilitate the ruminal N-utilisation (Jin et al., 2016). In contrast, HAB bacteria are known to massively increase NH_3 levels in the rumen fluid by their rapid AA deamination activity (Patra and Aschenbach, 2018). The subsequent steep increases of the ruminal NH_3 pool enhance NH_3 effluxes into the blood, increase the blood NPN-pool and promote the urea synthesis in the hepatocytes. Accordingly, HAB are thought to negatively influence the NUE and enhance N-losses (Hartinger et al., 2018).

Furthermore, this study identified three genera of *Lachnospiraceae* family – namely *Lachnospiraceae NK3A20 group*, *Lachnospiraceae AC2044 group* and *Lachnospiraceae XPB1014 group* – that possibly constituted microbial signatures influencing MU phenotypes. The *Lachnospiraceae* family has been associated with low N-utilising phenotypes in beef steers and goats (Wang et al., 2019, Alves et al., 2020). However, this family hosts a huge variety of genera that adapt to individual ecological niches and might therefore contribute to the N-metabolism in ruminants in a different manner (Meehan and Beiko, 2014). *Lachnospiraceae NK3A20 group* has been depicted as a major genus of the *Lachnospiraceae* family in the rumen (Anderson et al., 2021) and was significantly higher abundant in GBV_{HMU} cows. Interestingly, (Huang et al., 2021) stated a positive correlation between ruminal *Lachnospiraceae NK3A20 group* abundances and the rumen papillae length in yaks that quantitatively enhanced the absorption capacity of the tissue. Regarding higher occurrences of *Lachnospiraceae NK3A20 group* in cows with HMU phenotypes and correspondingly higher blood-N pools (Müller et al., 2021), it might be speculated if *Lachnospiraceae NK3A20 group* enhanced the NH_3 and NH^{4+} absorption by stimulating rumen papillae growth and thus promote N-effluxes into the blood. *Lachnospiraceae AC2044 group* and *Lachnospiraceae XPB1014 group* were higher abundant in GBV_{LMU} cows. Both genera were negatively correlated to MU_{lac}, and the *XPB1014 group* showed further considerable

negative correlations to MUY_{lac} and positive correlations to FaecN. Interestingly, *Lachnospiraceae XPB1014 group* abundances were found to be increased under N-scarcity in the hindgut of pigs and contributed to enhanced NUE (Zhao et al., 2020). Furthermore, increased *Lachnospiraceae XPB1014 group* abundance has been observed alongside with high carbohydrate fermentation levels in cows and pigs (Zhao et al., 2020, Hendawy et al., 2021). In general, a high fermentation rate of carbohydrates stimulates feed intake, accelerates the feed passage rate of the diet, reduces the absorption time of AA and peptides in the intestine, and thus increases faecal nutrient losses (McCarthy Jr et al., 1989, Höner, 2001, Brade and Distl, 2015, Schuba et al., 2017).
A further considerable positive correlation was found between FaecN and *Prevotellaceae UCG-003* abundances. Interestingly, high abundances of *Prevotellaceae UCG-003* were observed alongside with low ruminal fermentation degrees in yaks and steers (Liu et al., 2019, Qiu et al., 2020). Moreover, (Huang et al., 2021) identified a negative correlation between *Prevotellaceae UCG-003* and the length of the rumen papilla in yaks. Short papilla generally decrease the nutrient absorption capacity of the rumen tissue and might subsequently lead to enhanced faecal nutrient losses (Brade and Distl, 2015, Huang et al., 2021). The *Prevotellaceae* family is a major player in ruminal fermentation processes and hosts various genera, which are known to massively affect the ruminal AA metabolism (Liu et al., 2019, Qiu et al., 2020, Zhao et al., 2020). The present study revealed numerically higher abundances of *Prevotellaceae UCG-003* in GBV_{HMU} cows and further positive correlations to MU and MN. *Prevotellaceae UCG-003* might therefore constitute a microbial signature of GBVMU selection, which is proposed for independent effect estimation on faecal and milk N-losses from dairy cows.

Furthermore, *Anaerovibrio* and *Blautia* abundances were positively correlated to FaecN, MU and MN. Although, both genera were – in accordance with other studies – only barely abundant (Ramos et al., 2018, Xie et al., 2022) , *Anaerovibrio* and *Blautia* have been identified as differentially abundant in the digestive tract of goats grouped for high and low NUE phenotypes (Wang et al., 2019). *Anaerovibrio* was found to be significantly higher abundant in the low NUE goats, whereas *Blautia* was detected with higher presence in the high efficient phenotype. Although these findings are in accordance with significantly higher abundance of *Blautia* in GBV_{LMU} cows in the present study, the positive correlation between *Blautia* abundances and MU, FaecN and MN excretion would attribute *Blautia* a disadvantageous role in the ruminal N-metabolism. Since *Blautia* has just recently been defined as a single genus (Liu et al., 2021), further research on the biological contribution of *Blautia* to the N-metabolism in ruminants is proposed.

Moreover, the considerable positive correlation between *Butyrivibrio* abundances and FaecN attracted interest. *Butyrivibrio* is a dominant genus in the rumen (Henderson et al., 2015) and accounted for considerable abundances in the present study. *Butyrivibrio* has been identified with significantly higher occurrence in predisposed HMU cows in our previous study (Honerlagen et al., 2022) and correlated with lower N-recycling efficiency phenotypes in beef cattle (Alves et al., 2020). However, in the present study, *Butyrivibrio* accounted for significantly more abundances in GBV_{LMU} cows and correlated accordingly negative with MU_{lac}. Considering the major occurrence of this genus as well as the findings of (Derakhshani et al., 2018) who suggested *Butyrivibrio* as a major fybrolitic rumen dweller in Holsteins with a major impact on ruminal fermentation, it might be speculated whether *Butyrivibrio* affects the dietary N digestibility by its fermentation activity and therefore promote FaecN losses. At this time point, the causality between host genetics, *Butyrivibrio* abundances and the N-utilisation and N-excretion remains unclear, but deserves further research.

Conclusion

The results of this study implied that selection for GBV_{LMU} cows would reduce MU and UU but would not affect FaecN. Although GBVMU selection would potentially not lead to fundamental changes in the rumen microbial composition, specific genera abundances distinguished between GBV_{LMU} and GBV_{HMU} cows. Considering their relationship to MU and further NUE associated traits, *Succinivibrionaceae UCG-002*, *Clostridia unclassified*, *Desulfovibrio*, the *Lachnospiraceae* family, *Prevotellaceae UCG-003* and *Butyrivibrio* are proposed as the most considerable microbial genera linked to GBVMU, possibly

influencing proxies of NUE in Holsteins. These genera are suggested for quantitative effect estimations, in order to determine their potential as microbial signatures for future breeding selection on enhanced NUE in dairy cows.

Ethics statement

Animal housing and sampling were in accordance to the guidelines of the German Animal Protection Law. All protocols were approved by the Institute's Animal Welfare Commission. The sampling trial was conducted in strict compliance with the German Animal Welfare Legislation, has been approved by the Ethics Committee of the federal state of Mecklenburg-Western Pomerania, Germany (Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei; LALLF M-V7221.3-2-019/19) and is in accordance with the ARRIVE guidelines.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

HH organized the sampling trial. HH, IA and HR collected the samples. HH and IA conducted the laboratory work of rumen fluid samples. NT performed 16S sequencing. HH performed and HR assisted statistical analyses, interpretation of the data and writing of the manuscript. SP supported data analysis. DS assisted the data interpretation. NR, BK and KW conceived and supervised the study. All authors reviewed the final manuscript and provided critical feedback.

Funding

The study has received funding from the Federal Ministry of Food and Agriculture as part of the BlueCow Project (FKZ: 281B101516).

Acknowledgments

The authors thank the entire team of Gut Dummerstorf GmbH, led by Lucas Pieper, for the excellent support in carrying out the trial. Furthermore, we thank Tanja Lenke, Astrid Schulz and Thomas Geick for the elaborate analysis of N-contents in faeces samples as well as Solvig Görs and Kirsten Kàrpàrti who conducted the analysis of N-fractions in urine samples. We further thank Adewunmi Omolade Omotoso and Corinna Gladbach for their devoting help in the sample collection during the trial and Angela Garve, Angelika Deike and Marlies Fuchs for their assistance in laboratory analyses.

Suppl. capture

Suppl. 1. Statistics of microbe – trait correlations (p.1: correlation coefficients; p.2: adjusted *p*-values (Benjamini-Hochberg).

References

Abdoun, K., F. Stumpff, and H. Martens. 2006. Ammonia and urea transport across the rumen epithelium: a review. Anim Health Res Rev 7(1-2):43-59.

Alves, K. L. G. C., Y. T. Granja-Salcedo, J. D. Messana, V. C. de Souza, M. J. G. Ganga, P. H. D. Colovate, L. T. Kishi, and T. T. Berchielli. 2020. Rumen bacterial diversity in relation to nitrogen retention in beef cattle. Anaerobe:102316.

Anderson, C. J., L. R. Koester, and S. Schmitz-Esser. 2021. Rumen epithelial communities share a core bacterial microbiota: a meta-analysis of 16S rRNA gene Illumina MiSeq sequencing datasets. Frontiers in microbiology 12:625400.

Arunvipas, P., J. A. VanLeeuwen, I. R. Dohoo, G. P. Keefe, S. A. Burton, and K. D. Lissemore. 2008. Relationships among milk urea-nitrogen, dietary parameters, and fecal nitrogen in commercial dairy herds. Canadian Journal of Veterinary Research 72(5):449.

Bach, A., S. Calsamiglia, and M. Stern. 2005. Nitrogen metabolism in the rumen. Journal of dairy science 88:E9-E21.

Beatson, P., S. Meier, N. 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(10):2164-2171.

Bento, C. B. P., A. C. de Azevedo, E. Detmann, and H. C. Mantovani. 2015. Biochemical and genetic diversity of carbohydrate-fermenting and obligate amino acid-fermenting hyper-ammonia-producing bacteria from Nellore steers fed tropical forages and supplemented with casein. BMC microbiology 15(1):1-15.

Bergen, W. G. 2021. Amino acids in beef cattle nutrition and production. Pages 29-42 in Amino Acids in Nutrition and Health. Springer.

Bobbo, T., M. Penasa, A. Rossoni, and M. Cassandro. 2020. Genetic aspects of milk urea nitrogen and new indicators of nitrogen efficiency in dairy cows. Journal of Dairy Science 103(10):9207-9212.

Brade, W. and O. Distl. 2015. Das ruminale Mikrobiom des Rindes (1. Teil). Berichte über Landwirtschaft-Zeitschrift für Agrarpolitik und Landwirtschaft.

Bristow, A. W., D. C. Whitehead, and J. E. Cockburn. 1992. Nitrogenous constituents in the urine of cattle, sheep and goats. J Sci Food Agric 59(3):387-394.

Bryant, M. 1970. Normal flora—rumen bacteria. Am J Clin Nutr 23(11):1440-1450.

Burgos, S. A., N. M. Embertson, Y. Zhao, F. M. Mitloehner, E. J. DePeters, and J. G. Fadel. 2010. Prediction of ammonia emission from dairy cattle manure based on milk urea nitrogen: relation of milk urea nitrogen to ammonia emissions. J Dairy Sci 93(6):2377-2386.

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(7):1184-1196.

De Mendiburu Delgado, F. 2009. Una herramienta de análisis estadístico para la investigación agrícola. in Facultad de Ingeniería Industrial y de Sistemas. Sección de Postgrado. Universidad Nacional de Ingeniería, Lima (Perú), Universidad Nacional de Ingeniería, Lima (Perú), Facultad de Ingeniería Industrial y de Sistemas.

Derakhshani, H., J. C. Plaizier, J. De Buck, H. W. Barkema, and E. Khafipour. 2018. Association of bovine major histocompatibility complex (BoLA) gene polymorphism with colostrum and milk microbiota of dairy cows during the first week of lactation. Microbiome 6(1):1-18.

Difford, G. F., D. R. Plichta, P. Løvendahl, J. Lassen, S. J. Noel, O. Højberg, A.-D. G. Wright, Z. Zhu, L. Kristensen, and H. B. Nielsen. 2018. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. PLoS genetics 14(10):e1007580.

Dijkstra, J., O. Oenema, and A. Bannink. 2011. Dietary strategies to reducing N excretion from cattle: implications for methane emissions. Current Opinion in Environmental Sustainability 3(5):414-422.

Gonda, H. L. and J. E. Lindberg. 1994. Evaluation of dietary nitrogen utilization in dairy cows based on urea concentrations in blood, urine and milk, and on urinary concentration of purine derivatives. Acta Agriculturae Scandinavica A-Animal Sciences 44(4):236-245.

Grelet, C., E. Froidmont, L. Foldager, M. Salavati, M. Hostens, C. P. Ferris, K. L. Ingvartsen, M. A. Crowe, M. T. Sorensen, and J. F. Pierna. 2020. Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. Journal of dairy science 103(5):4435-4445.

Guliński, P., E. Salamończyk, and K. Młynek. 2016. Improving nitrogen use efficiency of dairy cows in relation to urea in milk–a review. Anim. Sci. Pap. Rep 34(1):5-24.

Hartinger, T., N. Gresner, and K.-H. Südekum. 2018. Does intra-ruminal nitrogen recycling waste valuable resources? A review of major players and their manipulation. Journal of animal science and biotechnology 9(1):1-21.

Hendawy, A. O., S. Sugimura, K. Sato, M. M. Mansour, A. H. Abd El-Aziz, H. Samir, M. A. Islam, A. R. Bostami, A. S. Mandour, and A. Elfadadny. 2021. Effects of selenium supplementation on rumen microbiota, rumen fermentation, and apparent nutrient digestibility of ruminant animals: a review. Fermentation 8(1):4.

Henderson, G., F. Cox, S. Ganesh, A. Jonker, W. Young, and P. H. Janssen. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Scientific reports 5(1):1-15.

Höner, K. 2001. Einfluss von Silagen aus unterschiedlichen Maishybriden auf die Umsetzungen im Verdauungstrakt von Kühen. Hannover, Tierärztl. Hochsch., Diss., 2001.

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. Frontiers in Genetics:1241.

Honerlagen, H., H. Reyer, D. Segelke, C. B. Müller, M. C. Prahl, S. Ponsuksili, N. Trakooljul, N. Reinsch, B. Kuhla, and K. Wimmers. 2022. Ruminal background of predisposed milk urea (MU) concentration in Holsteins. Frontiers in Microbiology:3131.

Huang, C., C. Liang, F. Ge, P. Yan, X. GUO, P. Bao, X. Ma, X. Wu, M. Chu, and X. Yao. 2021. Microbiome and Metabolomics Reveal the Effects of Different Feeding Systems on the Growth and Ruminal Development of Yaks. Frontiers in Microbiology 12:1440.

Huber, J. and L. Kung Jr. 1981. Protein and nonprotein nitrogen utilization in dairy cattle. Journal of Dairy Science 64(6):1170-1195.

Huhtanen, P., E. Cabezas-Garcia, S. Krizsan, and K. Shingfield. 2015. Evaluation of between-cow variation in milk urea and rumen ammonia nitrogen concentrations and the association with nitrogen utilization and diet digestibility in lactating cows. Journal of Dairy Science 98(5):3182-3196.

Jahnel, R. E., D. Wittenburg, M. Mayer, H. Taeubert, and N. Reinsch. 2021. Estimation of genetic parameters of milk urea content. Züchtungskunde 93(2):157-168.

Jin, D., S. Zhao, P. Wang, N. Zheng, D. Bu, Y. Beckers, and J. Wang. 2016. Insights into abundant rumen ureolytic bacterial community using rumen simulation system. Frontiers in microbiology 7:1006.

Lavery, A. and C. P. Ferris. 2021. Proxy Measures and Novel Strategies for Estimating Nitrogen Utilisation Efficiency in Dairy Cattle. Animals 11(2):343.

Libera, K., M. Szumacher-Strabel, M. Vazirigohar, W. Zieliński, R. Lukow, K. Wysocka, P. Kołodziejski, D. Lechniak, Z. Varadyova, and A. K. Patra. 2021. Effects of feeding urea-treated triticale and oat grain mixtures on ruminal fermentation, microbial population, and milk production performance of midlactation dairy cows. Annals of Animal Science 21(3):1007-1025.

Liu, C., H. Wu, S. Liu, S. Chai, Q. Meng, and Z. Zhou. 2019. Dynamic alterations in yak rumen bacteria community and metabolome characteristics in response to feed type. Frontiers in microbiology 10:1116.

Liu, X., B. Mao, J. Gu, J. Wu, S. Cui, G. Wang, J. Zhao, H. Zhang, and W. Chen. 2021. Blautia—a new functional genus with potential probiotic properties? Gut microbes 13(1):1875796.

Loor, J. J., A. A. Elolimy, and J. C. McCann. 2016. Dietary impacts on rumen microbiota in beef and dairy production. Animal Frontiers 6(3):22-29.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15(12):1-21.

McCarthy Jr, R., T. Klusmeyer, J. Vicini, J. Clark, and D. Nelson. 1989. Effects of source of protein and carbohydrate on ruminal fermentation and passage of nutrients to the small intestine of lactating cows. Journal of Dairy Science 72(8):2002-2016.

McMurdie, P. J. and S. Holmes. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol 10(4):e1003531.

Meehan, C. J. and R. G. Beiko. 2014. A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. Genome biology and evolution 6(3):703-713.

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(5):2468-2479.

Müller, C. B. M., S. Görs, M. Derno, A. Tuchscherer, K. Wimmers, A. Zeyner, and B. Kuhla. 2021. Differences between Holstein dairy cows in renal clearance rate of urea affect milk urea concentration and the relationship between milk urea and urinary nitrogen excretion. Sci Total Environ 755(2):143-198.

Munyaneza, N., J. Niyukuri, and Y. E. Hachimi. 2017. Milk Urea Nitrogen as an Indicator of Nitrogen Metabolism Efficiency in Dairy Cows: A Review. Theriogenology Insight 7(3):145-159.

Nousiainen, J., K. Shingfield, and P. Huhtanen. 2004. Evaluation of milk urea nitrogen as a diagnostic of protein feeding. Journal of dairy science 87(2):386-398.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. O'hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. Package 'vegan'. Community ecology package, version 2(9):1-295.

Paster, B. J., J. B. Russell, C. Yang, J. Chow, C. R. Woese, and R. Tanner. 1993. Phylogeny of the ammonia-producing ruminal bacteria Peptostreptococcus anaerobius, Clostridium sticklandii, and Clostridium aminophilum sp. nov. International Journal of Systematic and Evolutionary Microbiology 43(1):107-110.

Patra, A. K. and J. R. Aschenbach. 2018. Ureases in the gastrointestinal tracts of ruminant and monogastric animals and their implication in urea-N/ammonia metabolism: a review. Journal of advanced research 13:39-50.

Pérez-Enciso, M., L. M. Zingaretti, Y. Ramayo-Caldas, and G. de Los Campos. 2021. Opportunities and limits of combining microbiome and genome data for complex trait prediction. Genetics Selection Evolution 53(1):1-20.

Powell, J., C. Gourley, C. Rotz, and D. M. Weaver. 2010. Nitrogen use efficiency: A potential performance indicator and policy tool for dairy farms. Environmental Science & Policy 13(3):217-228.

Qiu, Q., C. Gao, M. Aziz ur Rahman, B. Cao, and H. Su. 2020. Digestive ability, physiological characteristics, and rumen bacterial community of holstein finishing steers in response to three nutrient density diets as fattening phases advanced. Microorganisms 8(3):335.

Ramos, A. F., S. A. Terry, D. B. Holman, G. Breves, L. G. Pereira, A. G. Silva, and A. V. Chaves. 2018. Tucumã oil shifted ruminal fermentation, reducing methane production and altering the microbiome but decreased substrate digestibility within a RUSITEC fed a mixed hay–concentrate diet. Frontiers in microbiology 9:1647.

Reed, K., L. Moraes, D. Casper, and E. Kebreab. 2015. Predicting nitrogen excretion from cattle. Journal of dairy science 98(5):3025-3035.

Rohart, F., B. Gautier, A. Singh, and K.-A. Lê Cao. 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS computational biology 13(11):e1005752.

Rosendahl, J. 2015. Untersuchung über die Beteiligung von nicht-selektiven Kationenkanälen am Ammoniaktransport über das Pansenepithel des Rindes. in Institut für Veterinär-Physiologie des Fachbereichs Veterinärmedizin. Freie Universität Berlin, Freie Universität Berlin, Institut für Veterinär-Physiologie des Fachbereichs Veterinärmedizin.

Ruska, D. and D. Jonkus. 2014. Crude protein and non-protein nitrogen content in dairy cow milk. Proc. Latv. Univ. Agr 32(327):36-40.

Saborío-Montero, A., M. Gutiérrez-Rivas, A. López-García, A. García-Rodríguez, R. Atxaerandio, I. Goiri, J. A. Jiménez-Montero, and O. González-Recio. 2021. Holobiont effect accounts for more methane emission variance than the additive and microbiome effects on dairy cattle. Livestock Science 250:104538.

Schuba, J., K.-H. Südekum, E. Pfeffer, and A. Jayanegara. 2017. Excretion of faecal, urinary urea and urinary non-urea nitrogen by four ruminant species as influenced by dietary nitrogen intake: A meta-analysis. Livestock Science 198:82-88.

Sørensen, P., M. Weisbjerg, and P. Lund. 2003. Dietary effects on the composition and plant utilization of nitrogen in dairy cattle manure. The Journal of Agricultural Science 141(1):79-91.

Spanghero, M. and Z. Kowalski. 2021. Updating analysis of nitrogen balance experiments in dairy cows. Journal of Dairy Science 104(7):7725-7737.

Stallcup, O., G. Davis, and L. Shields. 1975. Influence of dry matter and nitrogen intakes on fecal nitrogen losses in cattle. Journal of Dairy Science 58(9):1301-1307.

Tan, P., H. Liu, J. Zhao, X. Gu, X. Wei, X. Zhang, N. Ma, L. J. Johnston, Y. Bai, and W. Zhang. 2021. Amino acids metabolism by rumen microorganisms: Nutrition and ecology strategies to reduce nitrogen emissions from the inside to the outside. Science of the Total Environment 800:149-596.

Te Pas, M. F., T. Veldkamp, Y. de Haas, A. Bannink, and E. D. Ellen. 2021. Adaptation of livestock to new diets using feed components without competition with human edible protein sources—a review of the possibilities and recommendations. Animals 11(8):2293.

Wang, L., K. Liu, Z. Wang, X. Bai, Q. Peng, and L. Jin. 2019. Bacterial community diversity associated with different utilization efficiencies of nitrogen in the gastrointestinal tract of goats. Frontiers in microbiology 10:239.

Wood, G., P. Boettcher, J. Jamrozik, G. Jansen, and D. Kelton. 2003. Estimation of genetic parameters for concentrations of milk urea nitrogen. J Dairy Sci 86(7):2462-2469.

Xie, Y., H. Sun, M. Xue, and J. Liu. 2022. Metagenomics reveals differences in microbial composition and metabolic functions in the rumen of dairy cows with different residual feed intake. Animal Microbiome 4(1):1-12.

Yáñez-Ruiz, D. R., L. Abecia, and C. J. Newbold. 2015. Manipulating rumen microbiome and fermentation through interventions during early life: a review. Frontiers in microbiology 6:1133.

Zhao, Y., G. Tian, D. Chen, P. Zheng, J. Yu, J. He, X. Mao, Z. Huang, Y. Luo, and J. Luo. 2020. Dietary protein levels and amino acid supplementation patterns alter the composition and functions of colonic microbiota in pigs. Animal Nutrition 6(2):143-151.

Table 1: Means and standard deviations of NUE-associated traits in the whole cow population (n=358) and in the extreme groups (GBV_{HMU} (n=30), GBV_{LMU} (n=29)), respectively.

NUE-associated trait	Acronym	Entire herd	GBV_{HMU}	GBV_{LMU}	*p*-value[2]
Milk urea (mg/l)[1]	MU_{lac}	192.27 ± 24.80	226.97 ± 19.88	161.54 ± 19.10	<0.001
Milk urea (mg/l)	MU	174.04 ± 40.40	204.07 ± 45.42	146.73 ± 29.11	<0.001
Milk urea yield (g/d)[1]	MUY_{lac}	6.82 ± 1.36	7.87 ± 1.31	6.13 ± 1.38	<0.001
Milk nitrogen (g CP/100g milk)	MN	3.62 ± 0.38	3.64 ± 0.43	3.60 ± 0.40	0.355
Urinary urea (mmol/l)	UU	103.02 ± 50.52	131.12 ± 59.90	81.07 ± 54.69	<0.001
Faecal N (% in DM)	FaecN	2.50 ± 0.30	2.53 ± 0.27	2.49 ± 0.42	0.335
Milk yield (l/d)[1]	MY_{lac}	35.40 ± 5.68	34.62 ± 6.04	37.85 ± 6.06	0.022
Milk protein (% in milk)	MP	3.61 ± 0.38	3.62 ± 0.40	3.59 ± 0.40	0.403
Milk protein yield (kg/d)[1]	MPY_{lac}	9.87 ± 2.49	10.05 ± 2.15	9.98 ± 2.86	0.909

[1]lac = averaged data of one lactation (14 months);

[2]*p*-value < 0.05 indicates significant differences between GBV_{HMU} and GBV_{LMU} cow groups

Table 2: Correlations of NUE-associated traits in the whole cow population (n=358).

Trait	MU[1]	MU	MUY[1]	MN	UU	FaecN	MY[1]	MP[1]
MU[1]								
MU	0.63*							
MUY[1]	0.60*	0.31*						
MN	0.21*	0.25*	-0.11*					
UU	0.25*	0.26*	0.09	0.14*				
FaecN	0.06	0.10	0.03	-0.03	-0.07			
MY[1]	-0.06	-0.09	0.74*	-0.34*	-0.11*	-0.01		
MP	0.19*	0.28*	-0.12*	0.96*	0.15*	-0.02	-0.33*	
MPY[1]	0.25*	0.07	0.68*	0.06	0.06	-0.08	0.63*	0.07

[1]lac = averaged data of one lactation (14 months);

*indicates significant correlation (*p*-value < 0.05)

Table 3: Significantly differentially abundant genera (DAG) between the rumen fluids of GBV_{HMU} and GBV_{LMU} cows.

	Relative abundance[1] (%)				
Genus	Entire herd	GBV_{HMU}	GBV_{LMU}	Fold change[2]	*p*-value[3]
Anaerolineae unclassified[a]	0.0696	0.0943	0.0524	1.65	0.022
Blautia	0.0344	0.0243	0.0337	-2.15	0.008
Butyrivibrio	0.3634	0.3037	0.3841	-1.36	0.018
Clostridia unclassified	0.3234	0.3579	0.2937	1.38	0.044
Desulfobulbus	0.0424	0.0332	0.0438	-1.32	0.036
Desulfovibrio	0.0676	0.1113	0.0654	1.94	0.004
Lachnospiraceae AC2044 group	0.1293	0.1253	0.1628	-1.27	0.047
Lachnospiraceae NK3A20 group	2.5930	2.9075	2.2247	1.31	0.012
Lachnospiraceae XPB1014 group	0.0907	0.0758	0.0909	-1.32	0.041
Muribaculaceae ge	1.9545	2.1735	1.8869	1.14	0.015
Prevotellaceae NK3B31 group	0.3637	0.3367	0.4309	-1.31	0.013
Succinivibrionaceae UCG-002	0.2743	0.2731	0.2782	-2.68	0.043
Veillonellaceae unclassified	0.7186	0.9950	0.5936	2.48	0.043

[1]calculated as mean of the respective group;

[2]Fold changes derive from DESeq2 analysis (n=59 cows) and refer to GBV_{HMU} compared to GBV_{LMU};

[3]*p*-values < 0.05 indicate significance;

[a]Microbial genera underlined were also identified by sPLS-DA analysis to distinguish GBV_{HMU} and GBV_{LMU} rumen profiles

Table 4: Correlation between the abundances of potential microbial signatures derived from GBVMU selection with proxies of NUE in milk, urine and faeces in the entire cow population (n=358 cows)

Genus	MU_{lac}	MU	MUY_{lac}	MY_{lac}	MN	UU	FaecN	MP	MPY_{lac}
Anaerolineae unclassified[a]	-0.09	-0.05	-0.14*	-0.10	0.03	-0.14	0.09	0.01	-0.21*
Anaerovibrio	0.00	0.16*	-0.05	-0.08	0.16*	0.07	0.25*	0.17*	-0.03
Anaerovoracaceae ge	0.00	-0.14	0.12	0.17*	-0.11	0.04	-0.20*	-0.11	0.13
Bacteroidales BS11 gut group ge	-0.07	0.07	-0.23*	-0.23*	0.03	-0.11	0.22*	0.04	-0.30*
Blautia	-0.03	0.15*	-0.11	-0.13	0.20*	0.02	0.29*	0.21*	0.00
Butyrivibrio	-0.15*	0.06	-0.13	-0.06	0.07	0.01	0.24*	0.09	-0.17*
CAG-352	-0.11	0.06	-0.21*	-0.20*	0.08	-0.08	0.23*	0.10	-0.30*
Clostridia UCG-014 ge	-0.05	-0.11	0.04	0.12	-0.08	-0.03	-0.08	-0.08	0.01
Clostridia unclassified	0.00	-0.18*	0.04	0.07	-0.17*	0.08	-0.21*	-0.16*	0.07
Desulfobulbus	-0.04	0.03	0.08	0.14	0.00	0.04	0.00	0.00	0.09
Desulfovibrio	0.02	-0.06	0.04	0.05	-0.14*	-0.02	-0.16*	-0.17*	-0.05
Erysipelotrichaceae UCG-002	-0.03	-0.10	0.05	0.09	-0.07	0.06	-0.21*	-0.08	0.10
Erysipelotrichaceae UCG-007	-0.02	-0.17*	0.02	0.06	-0.08	0.06	-0.21*	-0.09	0.10
Firmicutes unclassified	-0.03	-0.03	-0.12	-0.12	0.03	-0.08	0.03	0.03	-0.22*
Lachnospiraceae AC2044 group	-0.13	-0.10	0.00	0.12	-0.11	-0.05	-0.07	-0.10	-0.04
Lachnospiraceae NK3A20 group	0.13	0.05	0.09	-0.01	0.15*	0.01	-0.07	0.15*	0.10
Lachnospiraceae XPB1014 group	-0.15*	0.03	-0.14*	-0.06	0.03	-0.09	0.21*	0.05	-0.21*
Muribaculaceae ge	-0.01	-0.03	-0.14	-0.17*	0.02	-0.05	0.21*	0.01	-0.20*
Prevotellaceae NK3B31 group	-0.16*	-0.06	-0.08	0.02	0.00	-0.08	0.12	0.00	-0.05
Prevotellaceae UCG-003	0.00	0.15*	-0.10	-0.13	0.10	-0.04	0.24*	0.11	-0.11
Succinivibrionaceae UCG-002	-0.06	0.10	-0.06	-0.05	0.05	0.00	0.11	0.07	-0.04
Syntrophococcus	0.03	-0.16*	0.11	0.12	-0.12	0.08	-0.22*	-0.12	0.14
Veillonellaceae unclassified	0.07	-0.10	0.09	0.09	-0.15*	0.07	-0.25*	-0.17*	0.09
WCHB1-41 ge	-0.04	0.13	-0.18*	-0.21*	0.12	-0.07	0.23*	0.12	-0.18*

[a]Microbial genera underlined were identified by both, 'DeSeq2' and 'SPLS-DA' analyses;

*****indicates significant correlations (adj. *p*-value<0.05)

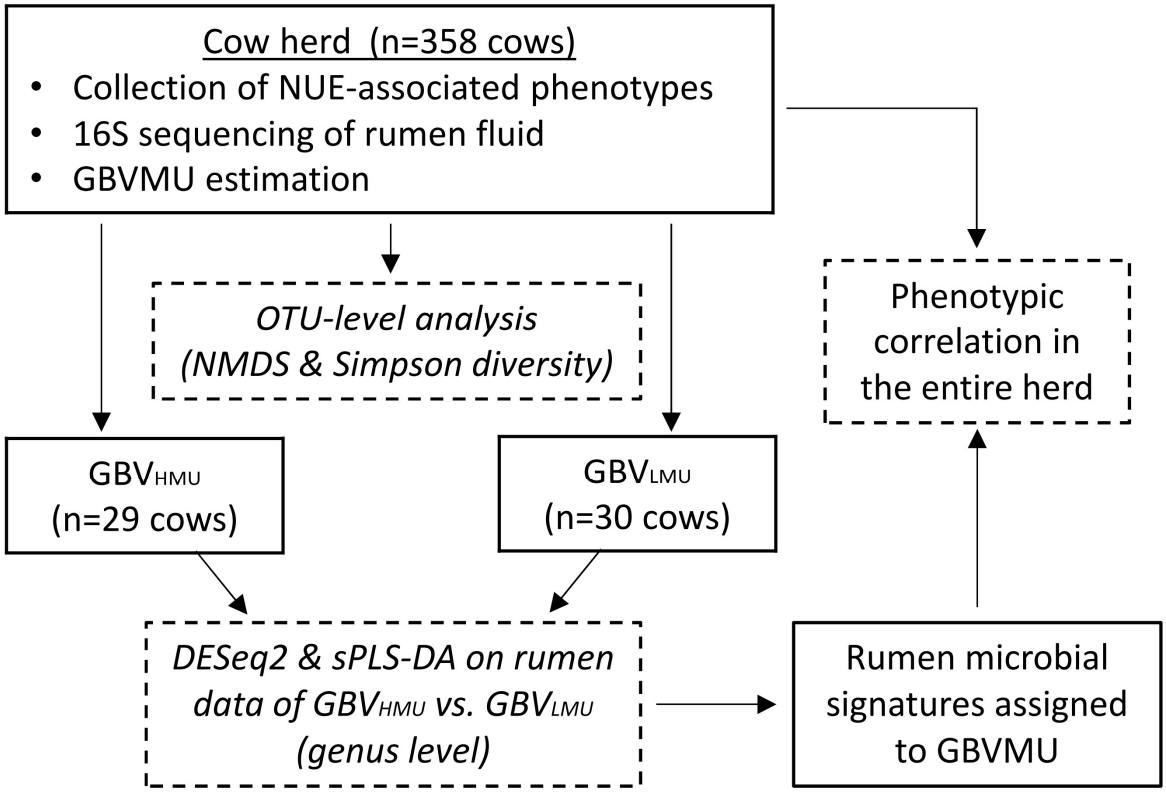

Figure 1. Flow chart of data collection and analysis.

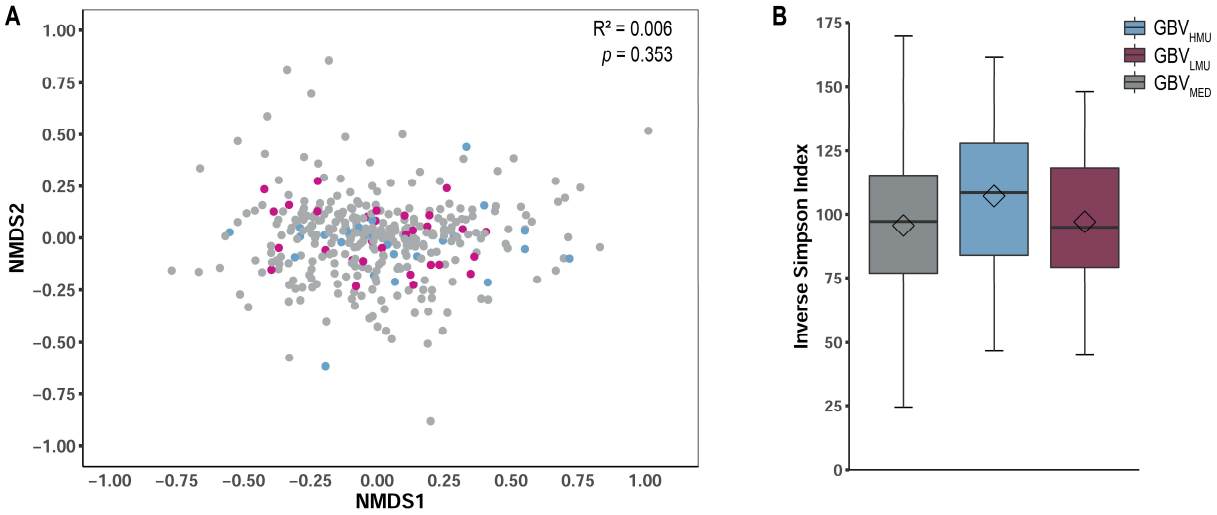

Figure 2. Analyses at OTU level did not indicate distinct clusters for GBVMU groups (A) nor did the alpha diversity significantly differ between the groups (B).

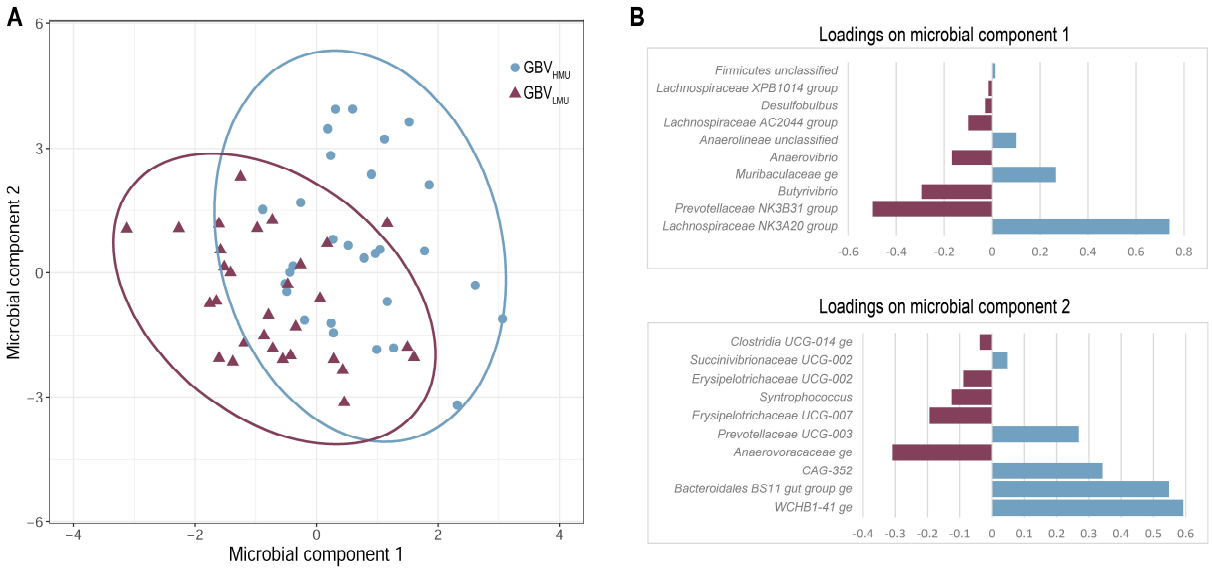

Figure 3. Sparse Partial Least Square-Discriminant Analysis (sPLS-DA) in dependence on extreme GBVMU groups, generated by the loading vectors (i.e. selected microbial genera) of component 1 and component 2.The ellipses (A) account for 0.95 confidence interval. The bar lengths of the microbial vectors (B, C) correspond to the importance of the respective microbial feature driving the respective component. The colour code of the vectors correspond to the cow group with higher median value of the respective microbe.

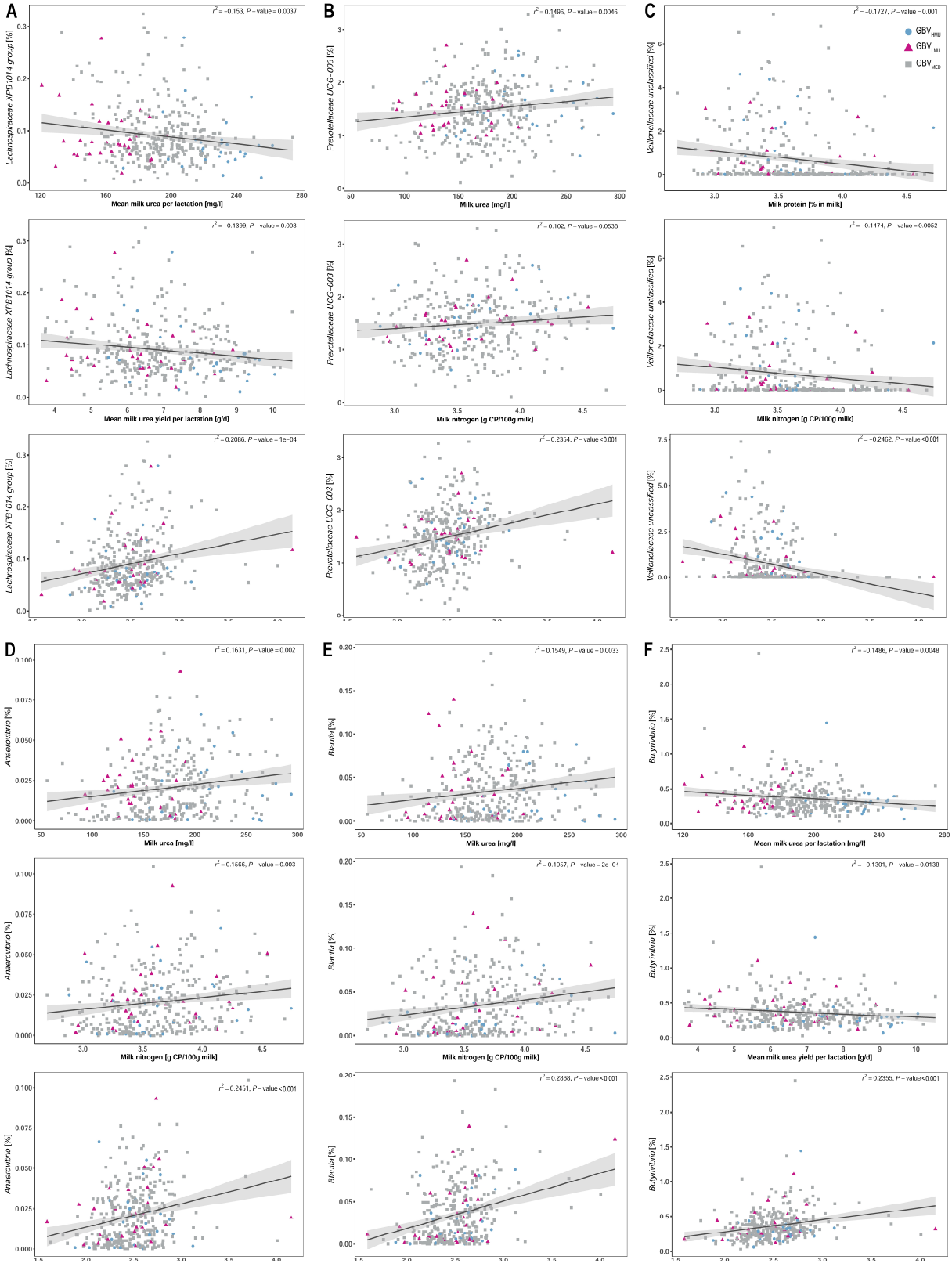

Figure 4A-F. Most prominent relationships between rumen microbial signatures assigned to GBVMU and specific proxies of NUE identified in the entire cow population (n=358). Dots display the microbial abundance – trait correlation coloured in correspondence to the respective GBVMU group (GBV_{HMU} (blue), GBV_{MED} (grey), GBV_{LMU} (red)) of the individual cows. The grey area displays the confidence interval (0.95).

11.5 List of scientific publications (peer-review) and presentations

I declare that I am the first author of all publications and presentations listed below. A detailed description of my contribution to the publications and experimental trials was described above.

List of Publications

Honerlagen, H., Reyer, H., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., Reinsch, N., and Wimmers, K. (2021). Identification of Genomic Regions Influencing N-Metabolism and N-Excretion in Lactating Holstein-Friesians. *Frontiers in Genetics*, 1241. doi: 10.3389/fgene.2021.69955

Honerlagen, H., Reyer, H., Segelke, D., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., and Wimmers, K. (2021). Rumen microbiota and host gene expression associate to predisposed milk urea (MU) concentration in Holsteins. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, (WCGALP), Volume Novel Traits, Wageningen Academic Publishers, 2022.

Honerlagen, H., Reyer, H., Segelke, D., Müller, C.B., Prahl, M.C., Ponsuksili, S., Trakooljul, N., Reinsch, N., Kuhla, B., and Wimmers, K. (2021). Ruminal background of predisposed milk urea (MU) concentration in Holsteins. *Frontiers in Microbiology*, 13. doi: 10.3389/fmicb.2022.939711

Honerlagen, H., Reyer, H., Abou-Soliman, I., Segelke, D., Ponsuksili, S., Trakooljul, Reinsch, N., N., Kuhla, B., and Wimmers, K. (2021). Microbial signatures inferred from association with genomic breeding values for milk urea concentration (MU) and their relationship to proxies for N-utilisation efficiency (NUE) in Holstein cows. *Journal of Dairy Science,* Submitted 23rd October 2022.

List of Presentations

Honerlagen, H., Reyer, H., Segelke, D., Oster, M., Ponsuksili, S., Trakooljul, N., Kuhla, B., and Wimmers, K. Ruminal background of predisposed milk urea (MU) concentration in Holsteins. 12th World Congress on Genetics Applied to Livestock Production (WCGALP, 2022), Rotterdam, the Netherlands.

Honerlagen, H., Reyer, H., Segelke, D., Reinsch, N., Kuhla, B., and Wimmers, K. Determinanten der kuhindividuellen Stickstoffutilisierung und –ausscheidung. Annual conference of the German Society for Animal Production and the Society for Animal Sciences (DGfZ/GfT-Gemeinschaftstagung, 2022), Kiel, Germany.

Honerlagen, H., Jahnel, R. E., Müller, C. B. M., Kuhla, B., Wimmers, K., Reinsch, N. Biomarkers for ruminal and endogenous N-use-efficiency to reduce N-emissions (BlueCow). Conference for the innovation promotion of climate protection in agriculture by the German Federal Ministry of Food and Agriculture (KlimAgrar, 2021), Berlin, Germany.

Honerlagen, H. Die Kuh als Klimaretterin? Annual conference of a Network for science, business, politics and education (Rostock's 11, 2021), Rostock, Germany.

Honerlagen, H. Ressourceneffizienz beim Rind. Annual meeting of the Committee for Animal Husbandry and Technology from the German Agricultural Society (FBN, 2021) Dummerstorf, Germany.

Honerlagen, H. Exploring determinants influencing individual ruminal and endogenous nitrogen utilization in dairy cattle. Seminar meetings, workshops, committees in and in between the Institutes of the FBN (FBN, 2019-2022), Dummerstorf, Germany.

11.6 Acknowledgement (Danksagung)

Ich möchte mich bedanken…

…bei Herrn Prof. Dr. Wimmers, für das Vertrauen mir dieses Thema zu übertragen. Danke, dass ich mich selbstständig, frei, mit Neugier und Freude wissenschaftlich entwickeln durfte und Sie gleichzeitig immer genau dann Orientierung gaben, wenn es klemmte. Danke, dass Sie sich immer Zeit für mich, meine Fragen und Gedanken genommen haben, auf Ihr Wort immer Verlass ist und Sie mir wertfrei und klar Antworten zu der Einordnung von Grundlagenforschung in den praktischen, landwirtschaftlichen Kontext geben konnten. Herr Prof. Dr. Wimmers, ich hatte eine unglaublich lehrreiche, intensive, spannende und prägende Zeit am FBN, vielen Dank, dass ich so viel lernen, ausprobieren und für mein Leben mitnehmen durfte.

…bei dem in meinen Augen besten Wissenschaftler des gesamten Institutes, der vorgibt, er sei Biologe mit Schwein- und Huhn Affinität, in seiner (Frei)zeit aber Pansensonden repariert, Kühe streichelt und sich mit Melkroboter- Dokus fortbildet. Lieber Henry, danke, dass du mich drei Jahre mit vollem Einsatz begleitet, unterstützt, aufgebaut und ausgehalten hast. Danke, dass du mit mir Tag und Nacht im Kuhstall standest, Wiederkäuer beprobt und Pansensäfte abgefüllt, Statistik erklärt, R Skripte und Endnote Libraries repariert, mit mir gepapert und gereviewed, dir am Telefon erst das linke und dann das rechte Ohr abkauen lassen hast und trotzdem (meistens) noch den Hörer abnimmst, du dich auch zwischenmenschlich so sehr um deine Doktorand (inn) en bemühst…und und und. Danke für viele tolle, bereichernde, wissenschaftliche Gespräche, für geduldige Antworten zu unendlich vielen Fragen, für deine prägnante Logik und deine ansteckende, wissenschaftliche Neugier. Ich kann mit Worten nicht ausdrücken, wie dankbar ich dir für die letzten drei Jahre bin und was für ein wissenschaftliches Vorbild du für mich bist. Danke, dass du mir geholfen hast mich zu entwickeln und mir so viel von dem wissenschaftlichen Handwerk beigebracht hast. Du bist (m)ein wissenschaftliches Vorbild.

…bei Lucas Pieper, Uwe August, Heidi und allen Melker (innen), die im Frühjahr 2020 auf Gut Dummerstorf tätig waren. Nur durch eure unproblematische, hilfsbereite, vertrauensvolle, aufmunternde und humorvolle Zuarbeit war die Beprobung der fast 400 Kühe und damit die Datengrundlage meiner wissenschaftlichen Arbeit möglich. Danke für einige der schönsten, aufregendsten, herzlichsten und intensivsten Tage und Nächte meines Lebens.

… bei jedem aus der Arbeitsgruppe (und dem erweiterten Kreis), die immer Anteil genommen haben, mich bei der Kuhbeprobung unterstützt, Proben eingefroren, mich im Labor angelernt, im Paperbusiness aufgebaut und bei Vorträgen mitgefiebert haben: Vielen Dank: Angela, Micha, Eduard, Tonn, Angelika, Ibrahim, Ade, Corinna, Fabio, Christian, Shahaf, Wietje, Ella, Frieder, Maruf, Hesi, Ma, Li, Maru, Kevin.

…bei Herrn Prof. Dr. Kuhla und bei Herrn Prof. Dr. Reinsch für telefonischen Auskünfte bzgl. meiner Fragen in der Ernährungsphysiologie und der quantitativen Genetik sowie bei Wencke Friedrichs und Silke Heinz als Ansprechpartnerinnen des MRV und bei Dierk Segelke (Vit Verden).

Ein unendlich großer Dank geht an meine Familie. Danke, Mama und Papa, dafür, dass Ihr mich immer unterstützt, zu mir haltet, mich machen lasst und trotzdem immer da seid, wenn ich Hilfe, oder Austausch brauche. Danke, dass ihr mich auch heute noch auffangt, wenn ich mich verloren habe. Danke Manni und Fiti, dafür, dass ihr euch immer interessiert, mitfiebert und mich nie im Stich lasst. Danke, dass ihr für meinen Geburtstag bis nach Rostock fahrt und mir hier so wunderbare Zuhause-Momente geschenkt habt. Danke, dass ihr die beiden besten großen Schwestern der Welt seid, auf die ich mich immer verlassen kann und bin so froh, dass wir alle wissen, was „Nützt ja nix“ bedeutet.

Ich möchte mich bei all meinen Freunden bedanken, die mich durch die letzte drei Jahre und länger begleitet haben, mein Leben so sehr bereichern, mich inspirieren mit mir lachen und weinen, tanzen und trinken, philosophieren und dummes Zeug reden, Orientierung geben oder mit mir das Leben entdecken. Danke, danke, danke, dass ihr alle mein Leben so schön, bunt und lebenswert macht. Weil ihr mich schon so lange und / oder so intensiv in den letzten drei Jahren begleitet habt, möchte ich mich ganz besonders bei Judith, Antonia, Kathrin M.-N., A & A Schamborski, Berit, Robert Weber, Alex, Ricarda und Schorlenjanni bedanken. Außerdem geht ein sehr großes Dankeschön an die Fishies & Friends Gruppe: Danke für sehr viele lustige, tiefe, mitfiebernde und mitfühlende Gespräche sowie für allerlei Unternehmungen in Dummerstorf, Rostock, Warnemünde und Schwerin. Danke, dass ich mit euch so eine tolle Zeit haben durfte: Alex, Tani, Dori und Marieke, ihr seid großartig!

11.7 Curriculum vitae

Hanne Honerlagen

Date of birth	20.02.1995
Nationality	German
Languages	German (native), English (good), French (basic)

Scientific Education

09/2019 – 2022	**PhD Student** Institute for Farm Animal Biology (FBN), Institute of Genome Biology, Dummerstorf (Germany) Thesis: Host genetic and rumen microbial determinants of nitrogen (N) utilisation and N excretion in lactating Holsteins
08/2019	**Master Degree in Livestock Science** Thesis: Pre-steps for the establishment of a genotyping assay with parentage control and simultaneous diagnostic of hereditary defects in dogs In cooperation with Generatio Sol. GmbH (69115 Heidelberg, Germany)
02/2018 – 08/2019	**Master in Livestock Science** Justus-Liebig-Universität Gießen (JLU), Germany
01/2018	**Bachelor Degree in Agriculture** Thesis: Fertility and milk yield – A relationship in dairy cows? Publication of the thesis by AV Akademikerverlag GmbH & Co. KG, Bahnhofstraße 28, 66111 Saarbrücken Title of publication: Milch oder Mutter – Die Kuh im Zwiespalt
09/2014 – 01/2018	**Bachelor in Agriculture** University of Applied Sciences, Kiel, Germany
2014	**Abitur (equivalent to A level)** Gymnasium Altenholz, Kiel, Germany