Elucidating biological aspects of feed efficiency in dairy cows using metabolomics

by

Mohsen Hashemiranjbar Sharifabad

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

ANIMAL SCIENCE

Department of Agricultural, Food and Nutritional Science

University of Alberta

Abstract:

Feed represents more than half of the costs of livestock production. Residual feed intake (RFI) is a phenotypic measure of feed efficiency that has been proposed as the best approach for genetic improvement of dairy cows. Although RFI is considered as a useful tool for feed efficiency, the application of RFI in dairy cows has been slow. The high cost of recording individual feed intake and production traits such as milk yield, body weight and milk composition traits has been the main reason for this slow rate of application. Keeping that in mind, the overall objectives of this study were: 1) identification of differences in metabolism of high and low RFI cows in early, mid, and late lactation stages and 2) assessment of the potential of metabolites as biomarkers for prediction of feed efficiency in lactating cows.

For RFI estimation of 75 lactating cows, feed intake and milk yield data were recorded on a daily basis, the concentrations of milk fat, protein and lactose were measured weekly and animals were weighed monthly. A random regression model (RRM) was used for prediction of daily values of component traits of RFI. Multiple linear regression was used to adjust actual energy intake for maintenance and production requirements. Values of RFI were obtained for each individual for day 3-240 of lactation and cows were grouped as high RFI (most inefficient, > +0.5 SD) and low RFI (least efficient, < -0.5 SD). The changes of RFI and its component traits for the RFI groups were visualized along lactation. RFI and actual energy intake (AEI) were significantly ($P<0.05$) different along lactation between groups. Whether these differences in feed efficiency have a basis in metabolism is investigated in three sampling days that are shown on RFI plots.

Targeted quantitative metabolomics using nuclear magnetic resonance (NMR) was employed to investigate the differences in metabolite profile of high and low RFI groups. Serum samples were taken at 50, 150 and 240 days in milk (DIM) from 75 cows and only high and low RFI

samples were analyzed. Concentration of 4 (glycerol, urea, creatinine, dimethyl sulfone), 4 (creatinine, glycerol, L-ornithine, L-lysine), and 8 (glycerol, acetone, citric acid, 3-hydroxy butyric acid, choline, creatinine, glycine, formate) metabolites were significantly ($P<0.05$) different at 50, 150, and 240 DIM respectively. In addition, a small number of metabolites showed a tendency to significance ($P < 0.10$) at 150 (acetone) and 240 DIM (isoleucine, L-lysine, urea).

Biomarker discovery analysis was done using principal component analysis (PCA), partial least square discriminant analysis (PLS-DA) and receiver operating characteristic (ROC) curve. The result showed discrimination of high and low RFI in PCA and PLS-DA score plots. This analysis enabled us to identify metabolites with the most discriminatory power between high and low RFI groups. Area under the curve (AUC) in ROC plots represents how well the model is capable of distinguishing between classes where the closer the value to one, the better is the model in classification. The PLS-DA model had AUC of 0.936, 0.866, and 0.997 for 50, 150, and 240 DIM respectively. The metabolites were then used for multiple linear regression to estimate RFI for individual cows. The adjusted R-squared of the models were 0.62, 0.65, and 0.83 for metabolite profile of serum samples from early, mid, and late lactation. The relative contribution of each metabolite was estimated. The metabolites that are significantly different in RFI groups have biological roles in feed efficiency. They are involved in the citric acid and urea cycles that are associated with energy and protein turnover at cellular level. Glycerol and creatinine were differentially expressed consistently in early, mid, and late lactation stages. This may be an indication that these biological pathways explain variation between high and low RFI animals. This study provides evidence to support the potential use of serum metabolites as biomarkers of

RFI for both classification into high or low RFI and prediction of the efficiency of individual cows, however, further investigation is warranted for validation of the results.

Acknowledgements

I would like to express my deepest gratitude to my supervisory committee, Dr. Graham S. Plastow and Dr Dagnachew Hailemariam for their constant support and scientific advice. They provided the financial support and created the environment that I could grow my skills, develop research ideas and experience the professional approach for scientific investigation. I am so grateful to them for this opportunity.

I am truly grateful to everyone at Livestock Gentec for their assistance, encouragement, and most importantly, friendship. I would also like to thank the Faculty of Graduate Studies and Research, Department of Agriculture, Food and Nutritional Sciences, and the University of Alberta for providing me with financial assistance.

A special thank you goes to my parents. I am much lucky to have you both as my parents and thank God for giving me you. I thank my wife, Anahid Hosseini who was with me step by step of this journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADG: Average daily gain

AEI: Actual energy intake

AFC: Age at first calving

BW: Body weight

DE: Digestible energy

DIM: Days in milk

DMI: Dry matter intake

EBV: Estimated breeding value

EBW: Empty body weight

EBWC: Empty body weight change

EEI: Expected energy intake

FCR: Feed conversion ratio

FPR: False positive rate

GC-MS: Gas chromatography – mass spectrometry

GE: Gross energy

GEBV: Genomic estimated breeding value

GEE: Gross energy efficiency

GFE: Gross feed efficiency

GHG: Greenhouse gas

GWAS: Genome wide association study

ICP-MS: Inductively coupled plasma mass spectrometry

LC-MS: Liquid chromatography – mass spectrometry

LMDB: Livestock metabolome database

LOD: Limit of determination

LOOCV: Leave one out cross validation

LOQ: Limit of quantification

LSM: Least square means

LTE: Lifetime efficiency

MCDB: Milk composition data base

miRNA: micro ribonucleic acid

mRNA: messenger ribonucleic acid

NEFA: Non esterified fatty acids

NMR: Nuclear magnetic resonance

PCA: Principal component analysis

PLS-DA: Partial least square discriminant analysis

RFI: Residual feed intake

ROC curve: Receiver operating characteristic curve

SD: Standard deviation

TMR: Total Mixed Ration

TPR: True positive rate

VFA: Volatile fatty acids

VIP: Variable of importance projection

VLDL: Very low-density lipid

Chapter 1. General introduction and literature review

1.1 Introduction

In most livestock operations, feed cost is more than half of total cost (Kennedy et al 1993). Although there have been improvements in growing crops efficiently to reduce feed cost for animals, other factors have kept prices high. Population growth, limited arable lands, and biofuel application for crops were among the top reasons. They have kept the livestock feed cost on a considerable fluctuation with an increasing trend in recent decades (Hemme & Otte, 2010).

For instance, global price trends for dairy products like milk shows during 1981 to 2005, world milk price ranged between 10-25USD per ton. However, it increased rapidly in 2007 to more than 45USD per ton as consequence of growth in price of skimmed milk powder (SMP) and butter from 1000 to 4000USD per ton, in response to lack of milk availability relative to world demand. On the other hand, feed cost for dairy farms has increased considerably and in parallel to milk price trend. World supplies of grain have not kept up with the growing demand for milk and it has led to historical high prices. Grain demand is driven by the need for feed and fuel of an ever-growing world population. Besides, higher energy prices push up the price for feed and land according to a Food and Agriculture Organization (FAO) report (Hemme & Otte, 2010).

This trend is also the case for the dairy industry in Canada. According to the CDC (Canadian Dairy Commission, 2016) cost of production studies, the amount of money spent for feed in dairy farms has increased from $12.80 per hectolitre of milk in 2012 to $17.61 in 2017. Thus, it shows how necessary it is to find a tool to create more feed efficient dairy cows because the animal is the core of the milk production system. There is a need to improve the production system with a focus on the dairy cow to reduce feed purchase costs.

Maximizing net income is generally the economic objective of a dairy farm. This can be achieved by conversion of a greater share of the feed nutrients into milk while less nutrient waste is emitted to the environment (de Ondarza et al. 2017). Feed efficiency could be the answer to this question on how to reduce feed cost while production is maintained. Different measures have been defined for feed efficiency of cattle to be used in breeding programs. Residual Feed Intake (RFI) is one of these feed efficiency measures that is used for economically important livestock species. It is the difference between actual and expected feed intake and was first defined by Koch, et al. (1963). We considered this measure of feed efficiency in lactating dairy cows and its relationship with blood serum metabolite profile in chapters 2 and 3, respectively.

After a brief introduction, the focus of this review will be first on efforts to measure RFI indirectly, metabolism and physiological aspects of feed efficiency. Then, I will investigate metabolomics technology and its success in biomarker discovery, and how a metabolomics approach can explain metabolic differences between groups of animals. Environmental concerns that can be reduced by genetic improvement in feed efficiency of dairy cows are mentioned at the end.

1.2 Definition and the objective of RFI estimation

The purpose of RFI estimation is genetic improvement of feed efficiency in farm animals. Residual feed intake (RFI) is the difference between actual and expected energy intake. This difference can be negative or positive. If an animal consumes less energy than our expectation, RFI will have negative value and this animal is feed efficient. The average daily RFI value for an individual in Mcal/day unit can be between -7 to +7 Mcal/day in 305 day of lactation where the average of this value for population is close to zero (Manafiazar et al. 2013).

The actual energy intake is estimated based on actual daily measurement of feed intake while the expectation of energy intake is more complex. The expectation of energy intake depends on animal maintenance and production requirements. These include basal metabolism, growth and milk production energy requirements. Therefore, regular body weight, milk yield, and milk composition traits also need to be recorded. Appropriate statistical modeling of these continuous energy-related traits can lead us to RFI values (Manafiazar et al. 2013).

Genetic evaluation of animals for feed efficiency needs a large sample size to enable researchers to capture the genetic diversity of the trait in a population of animals. Large sample size is also needed in order to reduce the standard error of estimations of genetic merit of animals. This is required to gain more accuracy to usefully estimate genetic merit of animals. More accuracy of genetic merit of animals will make the result of selection of animal for feed efficiency more reliable (Miglior et al., 2017).

One reason for the slow progress in genetic improvement of dairy cows using RFI has been its moderate heritability of RFI. Heritability is a parameter estimate of population that can show how much of phenotypic variation of a trait is due to genetic variation. Estimation of heritability is essential for selection and genetic improvement of animals. This is true while collection of phenotypic information of RFI for heritability estimation has been too expensive and slow. Collaboration of a group of countries including USA, Canada, UK and Netherlands has pooled feed efficiency data for estimation of heritability of RFI with good accuracy (Hardie et al. 2016). This study which was the outcome of multiple country efforts provided a suitably large sample size of RFI data to attain the desired accuracy. This shows the positive impact of large sample size on genetic improvement of RFI.

The high cost of data collection for estimation of RFI has been one of main barriers of genetic improvement of feed efficiency (VandeHaar et al. 2016). RFI is an expensive trait to measure and it takes a long time to collect the required feed intake and production data to sort animals based on RFI. Specifically, dairy cows have 305 days of milking which means it takes almost a year to measure feed efficiency accurately in dairy cows (Hardie et al. 2016). This has led many researchers to think about alternative methods for estimation of RFI in dairy and beef cattle.

1.3 Efforts to measure RFI indirectly

Indirect measurement of RFI using a method that is less expensive, fast, and reliably correlated with RFI is still a challenge. There are studies with different approaches to predict RFI indirectly. Application of infrared spectral profiles of milk samples as a predictor for DMI and RFI was considered recently (Shetty et al., 2017). The experiment was done using Danish Holstein and Jersey breeds. They developed models for indirect prediction of RFI. Cross validated R-squared was used as a measure of accuracy for how much of the variation in RFI could be explained by the model. They reported the cross validated R-squared of RFI for periods of early, mid, late and across lactation respectively to be 0.29, 0.09, 0.09 and 0.20 (Shetty et al., 2017). This showed that the model is really affected by peak of milk yield production period as early lactation RFI had the highest association with milk spectral data. Authors concluded that infra red spectral data does not have significant information in model development for prediction of DMI and RFI in comparison to the information that fat, protein and lactose provide for this purpose.

There are many more studies using this same experimental method with different analysis approaches, but none explained more variation of individual DMI of dairy cows (McParland et

al. 2014, Wallén et al. 2018). The common concept between them is that 1) inclusion of milk component traits and live body weight increases the explanatory power of models and 2) early lactation data would provide a better fit for prediction of DMI.

Another approach for prediction of RFI and DMI has been the use of thermal imaging to observe body surface temperature. Thermal imaging was used by Hardie et al. (2016) to measure surface body temperature to analyze its relationship with RFI and its potential use as an indicator for RFI. Although genetic analysis showed that surface temperature heritability was moderate the amount of variation that could be explained by body surface temperature was only 1.7%. Phenotypic correlation of RFI with surface temperature of paralumbar fossa, rump and leg were 0.12, 0.07 and 0.13. They suggested that the impact of environmental temperature on the quality of thermal images needed to be investigated in more detail in order to determine if this method could be useful (Hardie et al. 2016).

A similar study was done in beef cattle to predict RFI indirectly and had similar outcome. The correlation between infra red thermography (IRT) and RFI was significant (r=0.39) using infrared thermal cameras (Schaefer et al., 2018). However, an important condition of the procedure was that animals must be in their thermal neutral zone otherwise there is no significant correlation between RFI and body temperature. The experiment showed how environmental factors like wind and daily temperature can influence the neutrality of the thermal distribution on the body. They suggested further studies need to be done to consider important variables that affect this consideration (Schaefer et al., 2018).

1.4 Metabolism and physiological aspects of feed efficiency

The physiological basis of RFI in dairy cows is multi-dimensional and complex. It has not been well studied yet, however in beef cows more work has been done in recent decades (Richardson and Herd 2004, Davis et al. 2014). One of the fundamental differences in RFI model development between beef and dairy cows is the energy requirement of animals. In general, growth and maintenance energy requirements of cattle needs to be considered for both beef and dairy. The key difference is in lactation energy requirement consideration for dairy cows (Manafiazar et al. 2013). It is the part of feed efficiency in production dairy cows that make it more expensive and time consuming than beef cattle.

Although beef and dairy cattle are bred for different purposes, they are from one species and it is expected that they will have more similarities in feed efficiency than with other mammals. The review paper done by Richardson and Herd (2004) characterized RFI in cross-bred Angus beef steers. They concluded that inter-animal variation in RFI is composed of protein turnover and tissue metabolism and stress (37%), physical activity (10%), digestibility (10%), heat increment of fermentation (9%), body composition (5%) and feeding pattern (2%) whilst 27% was unknown or unassignable. The highest share of RFI sources of variation was due to protein turnover and it brought the attention of researchers to investigate the underlying metabolism differences between high and low RFI groups of animals. Thus, studies with designs that compare the extreme high feed efficient and low feed efficient animals in terms of biological differences has become more common in recent years. Blood, liver, and muscle are among the tissues that had most interest between researchers for investigation of biological differences. Liver plays a significant role in lipid metabolism and homeostasis of lipid in animals. For instance, long chain fatty acids are oxidized in hepatic mitochondria and peroxisomes. Then very

low-density lipoprotein (VLDL) carries triglycerides in the blood stream that are end products of oxidation. Non-esterified fatty acids (NEFA) in blood flow which is created by lipid metabolism can be oxidized by hepatic mitochondria and peroxisomes too (Dann & Drackley, 2005). Liver functions as an organ that metabolizes fat tissue and could show biological differences between high and low RFI cows due to the considerable amount of body fat that lactating dairy cows lose and recover along lactation.

Genetic inheritance of complex traits cannot be explained by segregation of a single gene. They are influenced by environmental and genetic factors. Expression of complex traits is impacted by many genes with small effects. To investigate the association of genetic variants with a trait, genome-wide association study (GWAS) can be done (Solem et al. 2008, Kadarmideen et al. 2008). RFI is a complex trait that has made it difficult to find causal genes by GWAS (Khansefid et al., 2017). Causal genes are genetic loci that are directly responsible for regulation of a trait. The expression of genes is also valuable information in detection and validation of causal genes (Solem, 2008). Therefore, differentially expressed genes between high and low RFI groups plus biological pathway analysis of significant genes that work together could provide insights on the genetic and physiological background of RFI (Khansefid et al., 2017).

The relationship of gene expression of liver and blood of Holsteins with RFI values and genomic estimated breeding values (GEBV) was considered recently by Khansefid et al. (2017). The normalized gene read counts which is a form of RNA data for gene expression studies was used in this study. The positive regression sign in association of gene expression and RFI was considered as upregulation of gene expression in high RFI cows and down regulation for low RFI cows. They discovered 233 and 240 genes that were upregulated and downregulated respectively in animals with high RFI. The GEBVs that were calculated from a previous study,

had significant association with 526 genes (Pryce et al., 2015). The study also identified 39 biological pathways in proteolysis, energy metabolism, regulation of transcription, translation, the cell cycle, and apoptosis. This study also considered Angus breed and reported 28 genes where expression levels in liver were associated with RFI and had the same magnitude as in Holstein breed (Khansefid et al., 2017). But at the same time, 285 and 454 genes in liver tissue of Holstein and Angus respectively were associated with RFI that had no overlap. The authors concluded that although not many genes that had significant association with RFI were overlapping, the same biological processes were involved. These biological function of these genes were protein catabolic processes, macromolecule catabolic processes, proteolysis involved in cellular protein catabolism, regulation of transcription, and oxidation reduction (Khansefid et al., 2017).

Liver gene expression has also been studied in beef cattle with different RFI values. A recent transcriptomic study compared high and low RFI animals and showed that most differentially expressed genes in liver are involved in many cellular processes such as lipid, amino acid and carbohydrate metabolism and energy production (Wang et al. 2020). They used whole transcriptome data from liver tissue as it is metabolically very active organ which shares 18-26% of total body oxygen for its activities. Angus, Charolais and Kinsella composite breeds were considered for this study. Although they were all beef breeds, the majority (82-88%) of differentially expressed genes in liver were breed specific. This shows a huge difference in biological regulation of genes associated with RFI in beef breeds: more than half of the differentially expressed genes between high and low RFI animals within a breed are not common with other beef breeds. It indicates the different genetic background of RFI in breeds of beef cattle.

The latter study also considered micro RNA (miRNA) expression between high and low RFI animals. The miRNA functions in (messenger RNA) mRNA silencing and post transcriptional regulation of gene expression. If a group of genes have significant contribution to the variation of RFI, it is wise to investigate which genes can silence the effect of them which can be those that produce miRNAs. This would also help provide additional information on the relevant biological networks and increase our understanding from gene interactions that build the final phenotype. The result of analysis of miRNAs were consistent with mRNAs results, miRNAs that are associated with RFI were also breed specific. The study identified 39 microRNAs associated with RFI that could target 55-76% of the genes that were involved in molecular metabolism (Wang et al. 2020). According to this study which used Ingenuity Pathway Analysis (IPA software) libraries for biological network analysis, both coding and non-coding DNA regions that are associated with RFI are significantly different between breeds of beef cattle. This shows there could be significant genetic variation between beef breeds for feed efficiency (Wang et al. 2020).

RFI is the residual of dry matter intake (DMI). It is reported that the phenotypic correlation between RFI and DMI in beef cattle is 0.73 (Crowley et al., 2010). There could be part of the variation of RFI hidden in metabolites associated with DMI. Therefore, another approach to reduce costs associated with RFI estimation is indirect prediction of DMI. Urine metabolite concentrations were found to predict a considerable amount of variation in DMI in beef and dairy cows in a recent paper (Dórea, Danés, Zanton, & Armentano, 2017). Furthermore, metabolomic approach could provide insight on biological regulation of DMI. Recently, Dórea et al. (2017) showed in a meta-analysis with a sample size of thousands of cows from independent studies (more than 3600 and 900 dairy and beef cattle respectively) that DMI could be predicted

based on urinary metabolites. This study demonstrated how much potential there is for using metabolites to accurately predict individual feed intake based on metabolite concentrations. They used 62 published data from both dairy and beef cattle and developed twenty models with body weight and metabolite concentrations of urine. Allantoin, uric acid and creatinine were used to predict DMI and digestible DMI. R-squared were reported in the range of 0.58 – 0.84 for dairy cows and 0.32 – 0.83 for beef cattle depending on the included predictor and formula of the model. Purine derivative and creatinine index (PDC index) which was suggested first by Chen, et al. (2004) had the best result. PDC index is urinary purine derivative to creatinine ratio that has been used as a non-invasive estimation of microbial protein synthesis in the rumen (Pulido & Briones, 2004). PDC index explained 75 and 67 percent of variation of DMI in beef and dairy cattle when it was not adjusted for random effect of the study. They discovered a stronger linear relationship ($R^2 = 0.90$ for beef, $R^2 = 0.96$ for dairy) between PDC index and DMI if they included study as the random effect in the model in this meta-analysis.

However, the predictive ability of the prediction model based on urinary metabolites still needs to be validated by other researchers. If the model is accurate then there is a chance to reduce cost of obtaining individual feed intake for RFI calculation of cattle. The interesting point of this meta-analysis with large sample size was the consistent result of the model for both dairy and beef cattle using urinary purine derivative (PD) and creatinine excretion for model development. This indicates that part of the physiological basis of DMI is common in beef and dairy because the models used the same metabolites.

The physiological basis of RFI in dairy cow is not well known and more work needs to be done. So far, most of the published studies had a relatively small sample size (5-10 animals per group).

This leads to limitation of statistical power to detect the effects. Small effect predictors may have not been seen in studies with small sample size.

Ribonucleic acid molecules from genome such as messenger RNAs (mRNAs) transfer biological information from nucleus to ribosomes for protein synthesis. miRNAs are another product of genome that has post transcription role to prevent mRNAs to reach ribosomes. They regulate gene expression by targeting groups of mRNAs. These types of RNAs are only part of available transcriptomic data that control the cell metabolism. Proteins function as enzymes to make biological reactions efficient and in some cases possible. They also build bonds with carbohydrates and lipids and influence biochemical reactions considerably. Proteomic and metabolomic studies could draw part of the underlying biological map of the metabolism. In addition to phenotypic recording at farm level, different omics levels from genomic to transcriptomic, proteomic and metabolomic could show differences between high and low RFI groups in dairy cows. For example, Salleh et al. (2017) indicated 19 and 70 candidate genes using RNA-based transcriptomic profiling which are associated with regulation of feed efficiency pathways in Danish Jersey and Danish Holsteins. The differentially expressed genes ($P<0.05$) between low and high RFI groups were investigated by pathway analysis. They concluded that the primary immunodeficiency pathway has a significant ($FDR<0.001$) role in utilization of feed and the metabolism of lipids, sugars and proteins. Salleh et al. (2017) also discussed that down-regulation of primary immunodeficiency pathway in high RFI (less feed efficient) Holstein and Jersey cows suggests that low immunity may influence feed efficiency. The relationship between lactating dairy cow immunity and feed efficiency has been seen before. Arndt et al (2015) investigated the sources of variation of feed efficiency in Holstein between extreme high and low feed efficient lactating cows. They defined feed conversion efficiency as

milk production (kg/day) divided by DMI (kg/day) and grouped animals based on this definition of feed efficiency. They found that somatic cell count (SCC) is significantly (P-value = 0.03) less in high feed efficient cows than low feed efficient dairy cows (Arndt et al., 2015). They converted SCC to somatic cell score (SCS) and found that the difference in SCS explained 6.4% of the observed difference between milk production of two groups of animals. It was concluded that there is a possibility that immune response contributes to feed conversion efficiency in dairy cows (Arndt et al., 2015). Although they did not use RFI for this study, the statistical evidence (small P-values) that more immunity is reported for more feed efficient dairy cows from two independent studies is promising. Immunity is also another complex trait of interest in dairy cows for breeding and its role in regulation of feed efficiency needs more attention in future studies.

Moreover, Salleh et al. (2017) reported other biological pathways to be differentially expressed between high and low RFI dairy cows such as metabolism of xenobiotics, sphingolipid metabolism, arachidonic acid metabolism, retinol metabolism, starch and sucrose metabolism and ether lipid metabolism. These pathways are mainly related to metabolism of nutrients such as protein, fatty acids, and carbohydrates. Different studies show that biological basis of feed efficiency in beef and dairy cattle is not limited to a couple of physiological pathways and more studies need to be done.

Metabolic biomarkers for RFI in Holstein and Jersey breeds of dairy were investigated by Wang and Kadarmideen (2019). They found significant ($P<0.001$) differences in fatty acid group from plasma samples between animals from high and low RFI dairy cows. They also used transcriptomic data from a previous study for integrative pathway analysis (Salleh et al., 2017). Using only metabolomic data and metabolic and transcriptomic data both showed three

important pathways. Amino acyl tRNA biosynthesis, citrate cycle (TCA) and alanine, aspartate and glutamate metabolism were reported as key potential metabolic biomarkers from this study that needs to be validated. However, the study used only 10 Holstein and 10 Jersey cows, the differences between high and low RFI groups within breeds with only 5 animals in each group were observed and they concluded a larger sample size could show more significant metabolites as biomarkers of RFI.

As examples from the literature show in recent years, the complexity and dynamic nature of feed efficiency in the dairy cow has increased the interest in researchers to approach this trait with assistance of other omics. Metabolomics is closer than other omics such as genomics, transcriptomics, and proteomics to the external phenotype of animals. It was suggested that this approach will therefore open a door to many potential discoveries in functional genomics (Wishart 2005). Functional genomics is the attempt to describe gene functions and interactions. Metabolomics is an additional information to other layers of omics and which can help to provide a better explanation of biological pathways that are related to the trait of interest (Wishart 2005) – feed efficiency in this case.

Therefore it is proposed that the biological basis of RFI and the answer to the question of how genetic variation of animals can affect the phenotypic variation of RFI can be better understood with metabolomics. Metabolomics has been used to map the biological pathways associated with feed efficiency to expand the knowledge of underlying physiology of RFI in beef cattle (Karisa et al. 2014). Proteomics is also of interest as proteins have important enzymatic roles in feed efficiency (Fonseca et al., 2019). Simultaneous analysis of multi-omics data to map biological networks are expected to be used more in the future to improve our understanding of RFI and other complex traits (Zampieri, Vijayakumar, Yaneske, & Angione, 2019).

Future use of metabolite concentration in investigation of physiological aspect of RFI can be found with fluxomics. Flux is rate of biological reaction (metabolite concentrations of reactants and products) within a biological environment in an organism (Winter & Krömer, 2013). Genomic and metabolic data can be used in fluxomic studies for estimation of rate of changes of metabolic pathways between animals form divergent groups. Potential applications of fluxomic studies in livestock are discussed in a recent review (Goldansaz et al., 2017). This kind of study can be useful in understanding metabolic sinks and assessing nutrition and metabolic efficiency. Tissue development and transformation in each developmental stage of an animal is another reason that makes metabolite flux more interesting. Fluxomic studies are a next level of omics that can show physiological aspect of feed efficiency. Such study has not been done yet in cattle due to costs and difficulty of isotope labeling required for such studies (Goldansaz et al., 2017).

1.5 What is Metabolomics?

The metabolome can be defined as the complete set of low molecular-weight compounds (i.e. metabolites) produced by cells. These compounds include peptides, organic acids, carbohydrates, amino acids, lipids, nucleotides, vitamins, toxins, minerals, cofactors and any compound with a molecular weight < 2000 Da that is chemically produced or transformed during metabolism (Wishart et al. 2013). Originally, the metabolome was described in the context of metabolic control analysis (MCA) as the complete set of low molecular weight endogenous molecules synthesized by an organism. Eventually the definition of the metabolome was modified and redefined to mean the collection of all small molecules that can be measured within a biological system. This includes both endogenous and exogenous metabolites such as drugs, toxins, and pollutants. Using modern metabolomics techniques many attempts have been made to estimate

the total numbers of metabolites in a variety of organisms and to construct comprehensive metabolome databases. However, this has proved challenging as the size of the metabolome varies with the organism's diet or other factors such as age and depends on the organism being studied.

Environmental effects like ration quality have significant effects on the metabolome of different biofluids of dairy cows (Sun et al., 2015). Therefore, in response to the demand for detailed information on milk metabolites and facilitate further research into milk chemistry, a database with appropriate baseline for milk metabolite was recently provided (Foroutan et al., 2019). The Milk Composition Data Base (MCDB) which is freely available in electronic version has identified 2355 cow milk metabolites. The database has all the different names of a single metabolite, structure for both abundant metabolites with a concentration of more than 1μM and rare metabolites. The MCDB includes both water soluble and lipid soluble metabolites (Foroutan et al., 2019). Modern quantitative metabolomics techniques such as nuclear magnetic resonance (NMR), inductively coupled plasma−mass spectrometry (ICP−MS) and liquid chromatography−mass spectrometry (LC−MS) were applied in this study that led to quantification of 296 bovine milk metabolites from the 2355 identified metabolites. This data was aided by knowledge from literature and used to build MCDB database accessible online.

The rumen is the primary place for microbial fermentation and has a vital role in cattle, sheep, and goat. Chemical composition of rumen reflects the rumen microflora and diet. Digestion of cellulose-rich material is possible in the rumen. The rumen converts cellulose to a wide range of metabolites that are used by microbes for proliferation and replication. Some of these metabolites are absorbed by cattle as nutrients for maintenance requirement, muscle production and milk production (Saleem et al. 2013). The health of ruminants and especially high producing animals

like lactating Holsteins depends on efficient rumen metabolism. Almost half of dairy cows in a given herd are more likely to be exposed to rumen-associated or metabolic diseases (Bertram et al., 2009). Therefore, the bovine ruminal fluid metabolome database (BRDB) was established to enhance quantitative metabolomics in this field of research (Saleem et al. 2013). They combined direct flow injection, NMR, GC-MS and ICP-MS with computer aided literature mining for identification and quantification of metabolites in bovine ruminal fluid. The application of multiple platforms expanded the metabolome coverage. The baseline of 246 identified and quantified metabolites with concertation means and standard errors and related literature value for each metabolite has become available (Saleem et al. 2013). They concluded that ruminal fluid is a metabolically diverse biofluid with 28 different compound categories. It was not surprising as estimates show that just 1 mL of ruminal fluid has 1 million protozoa, 10–50 billion bacteria, and 1000s of yeast or fungi (Fiehn et al. 2005).

1.6 Metabolomics platforms

Analysis of the entire metabolome using a single chemistry-based method is impossible due to the complexity and heterogeneity of the components of the metabolome. Therefore, multiple platforms are used for whole metabolome analysis of fluids from an organism. Liquid Chromatography Mass Spectrometry (LC-MS) and Nuclear Magnetic Resonance (NMR) have been the most common platforms used in recent years for detection, identification and quantification of small molecule compounds of biological samples. Different analytical platforms have different advantages and disadvantages. Basic principles of NMR and LC-MS and how they are used for metabolomics will be briefly explained. There are other metabolomics platforms for analysis of samples such as immunodetection, capillary electrophoresis and

infrared spectroscopy. More in-depth introduction of above-mentioned metabolomics technologies were reviewed by Dunn & Ellis (2005).

Theory of NMR spectroscopy was first triggered by Wolfgang Pauli in 1924 and two years later a paper was published about independent measurement of NMR phenomena (Bloch, 1946). Bloch, Packard and Purcell were awarded the Nobel Prize in physics in 1952 for their discovery. Since then, characterization of different compounds using NMR has developed to be one of the main methods in chemistry and this platform was also developed for metabolomics studies. Chemical compounds are made of elements such as oxygen, hydrogen that are connected with chemical bonds. Each element has a specific number of protons and electrons. Electrons spin around protons and they can be affected by magnetic field. Total angular momentum of the nucleus of an element is called nuclear spin. When a chemical compound is exposed to a certain known external magnetic field, the nuclear spin will transfer from a low to a high energy level (Nielsen & Oliver, 2005). The basic principle of NMR based methods is that there are certain nuclei that have a spin that generates small, local, magnetic fields. When an external magnetic field is applied to a chemical compound (which consists of atoms and nuclei that are chemically bonded together), it is possible to transfer energy to the chemical's nuclei (or cause their nuclear spins to flip) from a low to a high-energy level by exposing the nuclei to a weak, oscillating magnetic field at a "resonance" frequency that matches each of the nuclei's spinning frequency (Griffiths et al., 2010). Chemical/nuclear composition, its chemical structure, the chemical environment and the strength of the external magnetic field determines the resonance frequency that causes a given nucleus in a chemical compound to spin flip or jump in energy. When a nucleus (or nuclei if one is referring to a molecule) returns to equilibrium (a low-level state) energy is released at a frequency that corresponds to the characteristic resonance frequency of

the nucleus (or nuclei) in the biological sample. This energy can be detected via a radio frequency receiver and the oscillating signal can be Fourier transformed to produce an NMR spectrum (Nielsen & Oliver, 2005). The important feature of the spectrum is its characterization by lines and the peak or signals at different positions that corresponds to each chemical shift (Griffiths et al., 2010).

The other feature that has made NMR a common platform for metabolimcs projects is the non-destructive aspect of this technique which does not need chromatography separation or steps of derivatization. Therefore *in vivo* studies on bacteria and metabolism monitoring have become possible (Dass et al., 2017).

Reproducibility of NMR is high, and it means the result of NMR-based analyses are consistent. High reproducibility means the metabolite concentration that is gained by the same methodology described in different papers are very similar. This fact has increased the interest in using NMR in recent years for metabolomics studies (Emwas et al. 2015).

Simultaneous quantification of several classes of metabolites is a capability of NMR. Metabolites from different chemical classes have their own unique features that can be detected by NMR (Vadla, et al. 2013). For example protein-bound metabolites like lipoprotein particles and also certain inorganic ions or metabolites like H+ and metal ions that cannot be accurately determined with LC-MS or GC-MS can be analyzed by NMR (Ciborowski et al., 2012).

A disadvantage of NMR is a relatively low sensitivity (1-5 μM), however, it can be improved with higher magnetic field strength, hyperpolarization methods and cryo-cooled and micro probes. In general NMR spectroscopy is often 10 to 100 times less sensitive than LC-MS and GC-MS (Emwas et al. 2015). NMR-based metabolomics studies usually cover 50 – 200 metabolites that are present at over 1 μM while LC-MS is able to return over 1000 metabolites

with concentrations from 10 – 100 nM (Bertram et al., 2009).

Mass spectrometry-based methods have been used widely for comprehensive profiling of metabolites which provides both qualitative and quantitative information from structure and concentration of molecules within a sample. MS needs ionization methods to increase the number of metabolites that are detected (Guo, Zhang, Elmore, & Vishwanathan, 2013). The advantage of MS over NMR is its higher sensitivity of detection at nano molar concentrations and far more detectable number of metabolites, however, the disadvantage is that its results have been less reproducible than NMR. For a quantitative analysis, NMR is inherently quantitative because intensity of signal is directly proportional to number of nuclei in the molecule and metabolite concentration. MS intensity data does not have correlation with metabolite concentration most of the time and the reason is the different efficiency of ionization of different molecules (Emwas et al. 2015).

Technical advantages and disadvantages of both platforms were discussed above. Another important variable in platform considerations is cost of metabolite quantification for each method. Using both platforms is recommended (Marshall and Powers, 2017), as the metabolome is not detectable using a single platform. Biomarker discovery needs high sensitivity which MS methods provide but as indicated MS is less reproducible than NMR. Our study of association of metabolites and RFI has used NMR as it is a better approach for quantitative analysis and has better reproducibility (Marshall and Powers, 2017).

1.7 Metabolomics in livestock species

Metabolomics studies in livestock have mainly focussed on animal health with the bovine metabolome having the highest number of research publications. After cattle with 50% of the

papers, pig and sheep had 28 % and 13 % respectively. Most research (65%) were about health, production, and nutrition (Goldansaz et al., 2017). However, this report of application of metabolomics in livestock is from 2017, the need for biomarkers that can contribute to diagnosis or prognosis of diseases in livestock is evident. For instance, the application of metabolomics in identification of predictive biomarkers for post-partum disease states in dairy cows was reported by Hailemariam et al. (2014). They could predict which cows would develop periparturient disease 4 weeks before clinical symptoms using metabolites as biomarkers from plasma samples. The prediction had 87% sensitivity and 85% specificity using metabolite concentrations. This experiment in dairy cows using metabolomic approach can be also done for other complex traits such as RFI. This is done lately for high and low RFI cow and could show some groups of metabolites are expressed differentially (Kadarmideen et al. 2019). Metabolites in the fatty acid group had different concentration ($P<0.001$), however more work needs to be done for biomarker discovery that can classify high and low RFI lactating cows.

One of the challenges in the application of metabolomics after biomarker discovery is estimation of heritability of the metabolite. For example, beta hydroxy butyrate (BHB) is an indicator of ketosis in dairy cows and has a heritability between 0.25 to 0.36 in different stages of lactation while ketosis has heritability of 0.078 (Belay, 2017). If the metabolite has a moderate to high heritability like BHB, then this metabolite is more likely a good stable candidate to estimate the susceptibility of an individual for disease in a population of dairy cows or the merit of animals in feed efficiency. Heritability is a measure that can show us whether heritable genetic basis of metabolite is high enough that we expect to see this biomarker work well in future generations that will have different environmental effects. Estimation of heritability of indicator traits or intermediate phenotype has been discussed at transcriptome, proteome and metabolome level of

biology (Kadarmideen 2008). When the biomarker has moderate to high heritability, there is a genetic background for expression of metabolite in the population that most likely would not much change in one generation. But this needs a large sample size to show genetic variability of metabolite concentration in a population of animals (Jones, 2009). The cost, time and technological infrastructure are challenges for estimation of heritability of metabolites.

Despite the cost and technological needs of estimation of heritability for metabolites, there have been different studies in search of metabolites that have moderate to high heritability and strong association with complex traits. For instance, strong association between postpartum negative energy balance length with fertility was the motivation for researchers to study metabolites of blood plasma samples (Hayhurst, et al 2009). The importance of glucose and fatty acids function in negative energy balance length was the reason to estimate their heritability in blood plasma samples taken from dairy lactating cows (Hayhurst et al., 2009). Heritability estimates for glucose and free fatty acid (FFA) were moderate in the range of 0.23 ± 0.11 and 0.25 ± 0.13 respectively. All metabolite heritabilities were significantly different from zero. The authors indicated that metabolite heritability estimates were comparable to the heritability of other studies that measured the reproductive trait of interest (Hayhurst, et al 2009).

In addition to blood, milk has been one of the biofluids of interest that can potentially provide biomarkers of health status (or other traits) in dairy cows (Sundekilde, 2012). Fatty liver syndrome, rumen acidosis, milk fever and ketosis have been among important diseases studied using metabolomics in dairy cows. For example 3-hydroxy-butanoic acid has been identified as a biomarker for ketosis (Wittenburg et al., 2013). Another important feature of milk is its nutritional value, for example bovine milk oligosaccharides in cow milk are very similar structurally to human milk oligosaccharides and have similar protective roles for the infant

intestine. These benefits of metabolites in milk and blood of dairy cow and the predictive power for some diseases has brought attention to the question of how much variability in the concentration of these important metabolites are due to genetic variability of the population? If they have moderate to strong heritability, there could be chances for metabolites to be introduced as a trait for genetic selection. Wittenburg et al. (2013) reported the narrow-sense heritability for metabolites in the range of zero to 0.569. The study used Single Nucleotide Polymorphism (SNP) data with genomic best linear unbiased prediction (GBLUP) approach for each metabolite as response variable. Assessment of predictive power of SNPs in the model to predict metabolite measurements was done using leave-one-out cross validation (LOOCV). They estimated population genetic parameters such as additive genetic effect, dominance and polygenetic variance components for metabolites. It was concluded that the prediction accuracy of metabolites using SNP is at a high level. Correlation between observed and predicted metabolites were reported as 0.665, 0.668, 0.700, 0.709 and 0.769 respectively for Glycerol-3-phosphate, Glucose-2-amino-deoxy, Lactic acid, Ribulose-5-phosphate and Glucose 1,6-anhydro β.

The application of the metabolite BHB in detection of ketosis has a longer history in dairy cows. It is used as a gold standard indicator of ketosis with threshold of 1.2-1.4 mmol/L. In addition to the contribution to our knowledge of ketosis development, it has application in prevention of the disease progress by change of ration or treatments. Furthermore, there is evidence that genetic selection using BHB as an indicator of ketosis could make more genetic improvement than direct genetic selection for ketosis (Belay et al. 2017). This was investigated by the Canadian Dairy Network (CDN) using a pedigree of over 83 thousand animals, milk BHB concentration and incident of ketosis records. They concluded that selection for animals with lower BHB concentration can be used to increase the resistance to ketosis (Koeck et al., 2016).

New metabolomics approach with platforms such as NMR that can measure concentration of metabolites has provided an opportunity to detect potential biomarkers. Metabolites that can be indicators of RFI or at least classify animals into high and low RFI with high accuracy in dairy cows may be useful in management and genetic improvement of RFI.

1.8 Overall objective

RFI is the difference between the actual and expected feed intake. Prediction of RFI is expensive because of the need for long term regular measurement of individual feed intake. Indirect measurement of RFI can reduce the cost and speed up phenotyping of more animals in a fixed time period. Based on previous studies in beef cattle, it is believed that more than half of variation in RFI originates from differences in metabolism (Herd & Arthur, 2009). This shows there is significant potential for investigation of metabolism of RFI in cattle. Metabolomics has also shown biological differences in RFI in dairy (Kadarmideen 2019); and beef (Karisa et al. 2014). Thus, in this project the metabolomics approach was used to investigate differences between high and low RFI in lactating dairy cows. The overall objective of the thesis is to identify serum candidate biomarkers that distinguish high and low RFI cows and predict individual RFI phenotypes from metabolites.

Estimation of residual feed intake in dairy cows requires recording daily feed intake, measuring body weight, collecting milk samples and composition analysis, sampling the diet and performing dry matter and nutrient composition analysis during the lactation period. Because of these requirements, collection of data for RFI estimation is costly and difficult in commercial farm conditions. Consequently, this limits the application of RFI to improve feed efficiency and reduce carbon footprints from dairy cows. To overcome this problem, blood metabolites can be

investigated as a continuous indicator trait of RFI or a biomarker that can classify animals into high and low RFI groups.

In this project, a comparative serum metabolomics study was undertaken between high and low RFI cows at day 50, 150 and 240 days in milk (DIM) during lactation for discovery of potential metabolite biomarkers. The three time points were selected taking into account the physiological dynamics during the lactation period and represent early, mid and late lactation stages.

We hypothesized that:

1) The high and low RFI cows differ in serum metabolite profiles at 50, 150 and 240 DIM.
2) Classification of high and low RFI cows is possible using metabolite profiles at 50, 150 and 240 DIM.
3) RFI values of individual cows can be predicted using metabolite concentrations at 50, 150 and 240 DIM.

1.9 References

Arndt, C., Powell, J. M., Aguerre, M. J., Crump, P. M., & Wattiaux, M. A. (2015). Feed conversion efficiency in dairy cows: Repeatability, variation in digestion and metabolism of energy and nitrogen, and ruminal methanogens. *Journal of Dairy Science*, *98*(6), 3938–3950. https://doi.org/10.3168/jds.2014-8449

Belay, T. K., Svendsen, M., Kowalski, Z. M., & Ådnøy, T. (2017). Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows. *Journal of Dairy Science*, *100*(8), 6298–6311. https://doi.org/10.3168/jds.2016-12458

Bertram, H. C., Kristensen, N. B., Vestergaard, M., Jensen, S. K., Sehested, J., Nielsen, N. C., &

Malmendal, A. (2009). Metabolic characterization of rumen epithelial tissue from dairy calves fed different starter diets using 1H NMR spectroscopy. *Livestock Science*, *120*(1–2), 127–134. https://doi.org/10.1016/j.livsci.2008.05.001

Bloch, H. (1946). The Nuclear induction Experiment. *Physical Review*, *70*(7).

Canadian Dairy Commission, C. D. (2016). Canadian Dairy Commission, (November). Retrieved from http://www.cdc-ccl.gc.ca/CDC/index-eng.php?id=3806

Chen, X. B., Jayasuriya, M. C. N., & Makkar, H. P. S. (2004). Measurement and Application of Purine Derivatives:Creatinine Ratio in Spot Urine Samples of Ruminants. In *Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives* (pp. 167–179). Springer Netherlands. https://doi.org/10.1007/978-1-4020-2844-1_20

Ciborowski, M., Lipska, A., Godzien, J., Ferrarini, A., Korsak, J., Radziwon, P., … Barbas, C. (2012). Combination of LC-MS- and GC-MS-based metabolomics to study the effect of ozonated autohemotherapy on human blood. *Journal of Proteome Research*, *11*(12), 6231–6241. https://doi.org/10.1021/pr3008946

Crowley, J. J., McGee, M., Kenny, D. A., Crews, D. H., Evans, R. D., & Berry, D. P. (2010). Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. *Journal of Animal Science*, *88*(3), 885–894. https://doi.org/10.2527/jas.2009-1852

Dann, H. M., & Drackley, J. K. (2005). Carnitine palmitoyltransferase I in liver of periparturient dairy cows: Effects of prepartum intake, postpartum induction of ketosis, and periparturient disorders. *Journal of Dairy Science*, *88*(11), 3851–3859. https://doi.org/10.3168/jds.S0022-0302(05)73070-8

Darrell D. Marshall and Robert Powers. (2017). Beyond the Paradigm: Combining Mass

Spectrometry and Nuclear. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Dass, R., Grudziąż, K., Ishikawa, T., Nowakowski, M., Dębowska, R., & Kazimierczuk, K. (2017). Fast 2D NMR Spectroscopy for In vivo Monitoring of Bacterial Metabolism in Complex Mixtures. *Frontiers in Microbiology*, *8*(JUL), 1306. https://doi.org/10.3389/fmicb.2017.01306

Davis, S. R., Macdonald, K. A., Waghorn, G. C., & Spelman, R. J. (2014). Residual feed intake of lactating Holstein-Friesian cows predicted from high-density genotypes and phenotyping of growing heifers. *Journal of Dairy Science*, *97*(3). https://doi.org/10.3168/jds.2013-7205

de Ondarza, M. B., & Tricarico, J. M. (2017). Review : Advantages and limitations of dairy efficiency measures and the effects of nutrition and feeding management interventions. *The Professional Animal Scientist*, *33*(4), 393–400. https://doi.org/10.15232/pas.2017-01624

Dórea, J. R. R., Danés, M. A. C., Zanton, G. I., & Armentano, L. E. (2017). Urinary purine derivatives as a tool to estimate dry matter intake in cattle: A meta-analysis. *Journal of Dairy Science*, *100*(11), 8977–8994. https://doi.org/10.3168/jds.2017-12908

Dunn, W. B., & Ellis, D. I. (2005). Metabolomics: Current analytical platforms and methodologies. *TrAC - Trends in Analytical Chemistry*, *24*(4), 285–294. https://doi.org/10.1016/j.trac.2004.11.021

Emwas, A. H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods in Molecular Biology*, *1277*, 161–193. https://doi.org/10.1007/978-1-4939-2377-9_13

Fiehn, O., Wohlgemuth, G., & Scholz, M. (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. In *Lecture Notes in*

Bioinformatics (Subseries of Lecture Notes in Computer Science) (Vol. 3615, pp. 224–239). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11530084_18

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., … Fukumasu, H. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-41813-x

Foroutan, A., Guo, A. C., Vazquez-Fresno, R., Lipfert, M., Zhang, L., Zheng, J., … Wishart, D. S. (2019). Chemical Composition of Commercial Cow's Milk. *Journal of Agricultural and Food Chemistry*, *67*(17), 4897–4914. https://doi.org/10.1021/acs.jafc.9b00204

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS ONE*, *12*(5), 1–26. https://doi.org/10.1371/journal.pone.0177675

Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., & Deigner, H.-P. (2010). Targeted Metabolomics for Biomarker Discovery. *Angewandte Chemie International Edition*, *49*(32), 5426–5445. https://doi.org/10.1002/anie.200905579

Guo, J., Zhang, M., Elmore, C. S., & Vishwanathan, K. (2013). An integrated strategy for in vivo metabolite profiling using high-resolution mass spectrometry based data processing techniques. *Analytica Chimica Acta*, *780*, 55–64. https://doi.org/10.1016/j.aca.2013.04.012

Hailemariam, D., Mandal, R., Saleem, F., Dunn, S. M., Wishart, D. S., & Ametaj, B. N. (2014). Identification of predictive biomarkers of disease state in transition dairy cows. *Journal of Dairy Science*, *97*(5), 2680–2693. https://doi.org/10.3168/jds.2013-6803

Hardie, L. C., Beitz, D. C., Fernando, R. L., & Krizan, Z. (2016). The genetic basis and improvement of feed efficiency in lactating Holstein dairy cattle. Retrieved from

http://search.proquest.com.ezproxy.library.uq.edu.au/docview/1872300140/

Hayhurst, C., Flint, A. P. F., Løvendahl, P., Woolliams, J. A., & Royal, M. D. (2009). Genetic variation of metabolite and hormone concentration in UK Holstein-Friesian calves and the genetic relationship with economically important traits. *Journal of Dairy Science*, *92*(8), 4001–4007. https://doi.org/10.3168/jds.2008-1130

Hayhurst, C., Sørensen, M. K., Royal, M. D., & Løvendahl, P. (2007). Metabolic regulation in Danish bull calves and the relationship to the fertility of their female offspring. *Journal of Dairy Science*, *90*(8), 3909–3916. https://doi.org/10.3168/jds.2006-731

Hemme, T., & Otte, J. (2010). Pro-Poor Livestock Policy Initiative Status and Prospects for Smallholder Milk Production A Global Perspective. *Food and Agriculture Organization of the United Nations*, 1–12.

Herd, R. M., & Arthur, P. F. (2009). Physiological basis for residual feed intake. *Journal of Animal Science*, *87*(14 Suppl), 64–71. https://doi.org/10.2527/jas.2008-1345

Jones, E. J. (2009). Using Heritability for Genetic Improvement. Retrieved October 10, 2020, from https://www.pubs.ext.vt.edu/404/404-084/404-084.html

Kadarmideen, H. N. (2008). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., … Plastow, G. S. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livestock Science*, *165*(1), 200–211. https://doi.org/10.1016/j.livsci.2014.03.002

Kennedy, B. W., van der Werf, J. H., & Meuwissen, T. H. (1993). Genetic and statistical

properties of residual feed intake. *Journal of Animal Science*, *71*(12), 3239–3250. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8294275

Khansefid, M., Millen, C. A., Chen, Y., Pryce, J. E., Chamberlain, A. J., Vander Jagt, C. J., … Goddard, M. E. (2017). Gene expression analysis of blood, liver, and muscle in cattle divergentlya selected for high and low residual feed intake. *Journal of Animal Science*, *95*(11), 4764–4775. https://doi.org/10.2527/jas2016.1320

Koch, R. M., Chambers, D., & Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *Journal of Animal Science*, *22*(2), 486–494.

Koeck, A., Jamrozik, J., Kistemaker, G. J., Schenkel, F. S., Moore, R. K., Lefebvre, D. M., … Miglior, F. (2016). Genetic and phenotypic associations of milk β-hydroxybutyrate with ketosis in Canadian Holsteins. *Canadian Journal of Animal Science*, *96*(3), 302–305. https://doi.org/10.1139/cjas-2015-0206

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., Basarab, J., Li, P., & Wang, Z. (2013). Prediction of residual feed intake for first-lactation dairy cows using orthogonal polynomial random regression. *Journal of Dairy Science*, *96*(12), 7991–8001. https://doi.org/10.3168/jds.2013-6560

McParland, S., Lewis, E., Kennedy, E., Moore, S. G., McCarthy, B., O'Donovan, M., … Berry, D. P. (2014). Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *Journal of Dairy Science*, *97*(9), 5863–5871. https://doi.org/10.3168/jds.2014-8214

Miglior, F., Fleming, A., Malchiodi, F., Brito, L. F., Martin, P., & Baes, C. F. (2017). A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science*, *100*(12), 10251–10271. https://doi.org/10.3168/jds.2017-

12968

Nielsen, J., & Oliver, S. (2005). The next wave in metabolome analysis. *Trends in Biotechnology*, *23*(11), 544–546. https://doi.org/10.1016/j.tibtech.2005.08.005

Pryce, J. E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W. J., Coffey, M. P., Hayes, B. J., & Goddard, M. E. (2015). Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. *Journal of Dairy Science*, *98*(10), 7340–7350. https://doi.org/10.3168/jds.2015-9621

Pulido, R. G., & Briones, M. (2004). Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives. *Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives*, (June 2018). https://doi.org/10.1007/978-1-4020-2844-1

Richardson, E. C., & Herd, R. M. B. (2004). Biological basis for variation in residual feed intake in beef cattle . 2 . Synthesis of results following divergent selection Cooperative Research Centre for Cattle and Beef Quality . *Australian Journal of Experimental Agriculture*, *44*, 431–440. https://doi.org/10.1071/EA02221

Saleem, F., Bouatra, S., Guo, A. C., Psychogios, N., Mandal, R., Dunn, S. M., … Wishart, D. S. (2013). The Bovine Ruminal Fluid Metabolome. *Metabolomics*, *9*(2), 360–378. https://doi.org/10.1007/s11306-012-0458-9

Salleh, M. S., Mazzoni, G., Höglund, J. K., Olijhoek, D. W., Lund, P., Løvendahl, P., & Kadarmideen, H. N. (2017). RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high- and low-residual feed intake in Nordic dairy cattle. *BMC Genomics*, *18*(1), 1–17. https://doi.org/10.1186/s12864-017-3622-9

Schaefer, A. L., Ominski, K., Thompson, S., Crow, G., Bench, C., Colyn, J., … von Gaza, H.

(2018). Energy utilization in cattle with steady state and non-steady state methods: the importance of thermal neutrality. *Heliyon*, *4*(10). https://doi.org/10.1016/j.heliyon.2018.e00843

Shetty, N., Løvendahl, P., Lund, M. S., & Buitenhuis, A. J. (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. https://doi.org/10.3168/jds.2016-11609

Solem, C., Koebmann, B., & Jensen, P. R. (2008). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Sun, H. Z., Wang, D. M., Wang, B., Wang, J. K., Liu, H. Y., Guan, L. L., & Liu, J. X. (2015). Metabolomics of four biofluids from dairy cows: Potential biomarkers for milk production and quality. *Journal of Proteome Research*, *14*(2), 1287–1298. https://doi.org/10.1021/pr501305g

Sundekilde, U. K. (2012). Milk metabolite variability and heritability and their association with technological properties of bovine milk elucidated by NMR-based metabonomics, (July), 228.

Vadla, N. C., Davalagar, V. D., & Sripadi, P. (2013). Detection and characterization of N-alkyl diethanolamines and N-2-alkoxyethyl diethanolamines in milk by electrospray ionization mass spectrometry. *Metabolomics*. https://doi.org/10.1007/s11306-012-0492-7

VandeHaar, M. J., Armentano, L. E., Weigel, K., Spurlock, D. M., Tempelman, R. J., & 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.

https://doi.org/10.3168/jds.2015-10352

Wallén, S. E., Prestløkken, E., Meuwissen, T. H. E., McParland, S., & Berry, D. P. (2018). Milk mid-infrared spectral data as a tool to predict feed intake in lactating Norwegian Red dairy cows. *Journal of Dairy Science*, *101*(7), 6232–6243. https://doi.org/10.3168/jds.2017-13874

Wang, X., & Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites*, *9*(7), 1–16. https://doi.org/10.3390/metabo9070151

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., … Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. *BMC Genomics*, *21*(1), 1–22. https://doi.org/10.1186/s12864-019-6362-1

Winter, G., & Krömer, J. O. (2013). Fluxomics - connecting 'omics analysis and phenotypes. *Environmental Microbiology*, *15*(7), 1901–1916. https://doi.org/10.1111/1462-2920.12064

Wishart, D. S. (2005). Metabolomics: The principles and potential applications to transplantation. *American Journal of Transplantation*, *5*(12), 2814–2820. https://doi.org/10.1111/j.1600-6143.2005.01119.x

Wishart, David S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., … Scalbert, A. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research*, *41*(D1), 801–807. https://doi.org/10.1093/nar/gks1065

Wittenburg, D., Melzer, N., Willmitzer, L., Lisec, J., Kesting, U., Reinsch, N., & Repsilber, D. (2013). Milk metabolites and their genetic variability. *Journal of Dairy Science*, *96*(4), 2557–2569. https://doi.org/10.3168/jds.2012-5635

Zampieri, G., Vijayakumar, S., Yaneske, E., & Angione, C. (2019). Machine and deep learning

meet genome-scale metabolic modeling. *PLoS Computational Biology*, *15*(7), 1–24. https://doi.org/10.1371/journal.pcbi.1007084

Chapter 2. Estimation of residual feed intake (RFI) for lactating dairy cows

2.1 Introduction

2.1.1 Measures of feed efficiency

There have been many studies on cattle feed efficiency exploring a number of different approaches. In recent years, there has been substantial work on the application of residual feed intake (RFI) as a good feed efficiency measure to describe variability of energy efficiency in dairy cows. However, several other methods have been proposed and they have different applications and limitations. The question of which definition of feed efficiency is better and suitable for genetic selection will be discussed here.

Feed conversion ratio (FCR) is commonly defined as kilogram concentrate fresh weight per kilogram product fresh weight. Feed conversion ratio is easy to use and more traditional common tool to calculate feed conversion in livestock. The aim is to reduce FCR in order to reduce the cost of production. FCR is generally used for nutritional studies or management purposes in monophagous animals and beef cattle. But it should be mentioned that sources of genetic variation of energy efficiency was basically reviewed by Veerkamp and Emmans (1995); they considered FCR for dairy cows. Manafiazar et al (2015) considered both FCR and RFI in dairy lactating cows and reported phenotypic and genetic correlation among them. However, later on scientists used Gross Feed Efficiency (GFE) for dairy cows because FCR did not have milk yield in its definition.

Gross Feed Efficiency (GFE) is the amount of milk yield divided by feed intake over a period of time (Wang and Kadarmideen 2019). Milk yield is the numerator and feed intake the denominator in this case. The higher the value of GFE, the more efficient the animal because she produces more milk for a fixed amount of feed. GFE can be easily calculated by farmers via

daily milk and DMI records. However, as for FCR, GFE is a ratio trait that brings a problem which will be discussed later.

The other index that has been often used in dairy cattle is Gross Energy Efficiency (GEE) which is the energy in milk divided by total energy intake (Veerkamp & Emmans, 1995). Unlike FCR, the higher GEE for a cow means a more feed efficient animal because output is the numerator of the ratio. GEE takes the energy of feed material into account and it is more reasonable in dairy cattle system as the moisture of feed has a large variation ranging from 60g/Kg in concentrate to 900g/Kg in some root crops (McDonald, Edwards, Greenhalgh, & Morgan, 2002). Rations for cattle diets are formulated with Dry Matter (DM) and it is highly correlated with Gross Energy (GE) of feed. The denominator for each feed ingredient can be calculated through multiplying dry matter (DM) of feed by a constant coefficient which comes from the gross energy of feed. Therefore, GEE has many advantages in comparison with FCR for dairy nutrition, but both of these plus GFE have drawbacks for animal breeding programs.

There are distinctive partial efficiencies for maintenance, pregnancy, lactation and body gain or loss and these two indices (FCR and GEE) are not able to demonstrate them separately (Veerkamp & Emmans, 1995). Furthermore, they have phenotypic and genetic correlation with production traits such as Average Daily Gain (ADG) which is a measure of growth and mature size. When an unfavoured significant correlation exists between GEE and ADG, selection for animals based on GEE would result in offspring with higher body size and consequently a greater maintenance requirement. Finally, we will not have any clue whether improvement has taken place from a change in the numerator or denominator of the GEE and FCR ratios. It is important for genetic improvement (or animal breeding) to know the direction of changes in a trait to be able to identify which components of efficiency of metabolism have the largest genetic

variation as this is the basis for improvement (Manafiazar, et al 2012). That is why ratio traits are problematic here.

Lifetime Efficiency (LTE) is another measure that takes the whole life of the cow into account. It is defined as "the capture of feed energy in milk, conceptus, and body tissue divided by gross energy intake during the life of cow, starting at birth" (Vandehaar, 1998; VandeHaar and St-Pierre, 2006). Many factors in LTE should be standardized to be able to compare one cow to others, such as age at calving, calving interval, feeding, and housing. It is a merit of this index that it looks at all economic life of a cow and production management to evaluate profitability of each cow. But it takes a long time to access this data and genetic progress of dairy industry is moving faster than the life cycle of a dairy cow. Dairy cows on average have an economic life of 4 years and it takes too long to measure the merit of this cow for selection because the merit of sires are based on the production data of daughters. Moreover, LTE is highly dependent on precalving and intercalving period because we considered energy efficiency of the animal for its whole life.

The other problem of all these ratio measures is that any phenotypic correlation between feed efficiency and other traits should be measured carefully as it can change the result of selection in a way which is not expected, and results of breeding programs can have potential effects for many generations.

Residual feed intake (RFI) was first suggested as a better alternative for beef cattle by Koch and colleagues (Koch, et. al 1963) and is defined as the difference between actual feed intake and the prediction of feed requirement based on the animal's weight and production. Koch (1963) proposed RFI for beef cattle with heritability of 0.28 using body weight and feed intake data from 1950s. He compared different scenarios of defining feed efficiency and showed the

importance of analysis of covariances between the definition of feed efficiency and feed consumption. This trait in dairy cattle is moderately heritable with a range of 0.22 to 0.38 (Korver et al., 1991; Pryce et al., 2014). The dairy cow undergoes different stages of physiological change during a lactation cycle and within this period feed efficiency changes. RFI does not have phenotypic correlation with BW, ADG and Energy Corrected Milk (ECM) yield (Connor et al. 2013; Manafiazar et al. 2015). Therefore, it would not have the same problems as indicated for ratio traits of feed efficiency. RFI also has a normal distribution that improves its utility for conventional genetic selection compared to other feed efficiency measures. RFI is an indicator of differences in metabolism rather than a direct interpreter of differences in production as Crews (2005) explained. Because RFI is not a ratio, it is equal to the difference of actual energy intake and expected energy intake. Positive RFI cows consume more than the expected energy requirement for maintenance and production while negative RFI cows consume less than the expected for the same production level. RFI calculation in research centers and universities using regular precise measurement of energy related phenotypes of individual animals is the first step to recognize feed efficient animals for genetic selection, however other aspects for RFI must be considered for genetic selection like phenotypic and genetic correlation of RFI with other economic traits. Negative RFI cows have lower feed intake compared the positive RFI cows for the same production level. Selection for negative RFI cows can reduce feed cost and improve profitability. The other benefit of selection for negative RFI cows is a reduction of enteric methane emission. The relationship between RFI and methane production in beef and dairy cattle was studied and reviewed by many authors (e.g. Arthur and Herd 2008, Basarab et al. 2013). Residual feed intake (RFI) is the difference between actual energy intake and expected energy intake. This measure of feed efficiency considers the maintenance, lactation, and growth energy

requirements of the cow during lactation. RFI is a useful feed efficiency measure as it can extract or identify the individual merit of an animal in terms of feed efficiency with high precision because it is going to be used for genetic selection which will have a long-term effect on the population.

Some researchers such as Berry and Crowley (2012) believe RFI may have lack of acceptance from producers. The reason is that in beef cattle, slow growing animals which eat relatively less feed may have low RFI. They proposed another feed efficiency measure, Residual intake and body weight gain (RIG) for only beef cattle. It has significant phenotypic and genetic correlation with production traits such as average daily gain (ADG). Positive values of their proposed measure show more feed efficient animals which was considered an advantage over RFI concept by authors. Animals with greater RFI that are more feed efficient have greater ADG and reduced DMI. To date, this measure of feed efficiency has not been redefined for dairy.

Different energy requirements are important during lactation. Maintenance, milk production and body reserve changes vary between individuals and we need a measure that can account for this variability (Manafiazar et al., 2012). The reason that we are interested to see variability of the trait is that phenotypic and genetic variation has a great importance for genetic selection programs. The first step for genetic improvement of RFI in dairy cow is to define the statistical model for RFI prediction. It has been done by Manafiazar et al. (2013) and this chapter will briefly explain the method and how it is used to separate animals into high and low RFI groups that can be used to discover metabolomic biomarkers as discussed in chapter three.

2.1.2 Estimation of RFI in lactating dairy cows

Residual feed intake (RFI) is a measure of feed efficiency that can explain individual phenotypic merit of individuals, however, its estimation needs complex statistical modeling in dairy cows. RFI is equal to actual energy intake (AEI) minus expected energy intake (EEI). In order to have a reasonably accurate estimate of RFI values for ranking a group of lactating dairy cows, multiple traits need to be observed over time. Major traits related to RFI such as dry matter intake, body weight, milk yield and milk composition traits. The complexity originates from infeasibility or high cost of daily recording of production traits such as animal body weight in every single day of lactation.

In practice, milk fat, protein, and lactose concentrations can be recorded weekly, body weight monthly, and feed intake on a daily basis in research facilities. These traits are used to build EEI which represent multifunctional energy requirements of lactating cows. Multifunctional energy requirements of dairy cows are maintenance requirement for basal metabolism, milk production energy requirement and body weight gain or loss along lactation.

Multifunctional energy requirements can be estimated as RFI components using metabolic body weight (MBW), milk production energy requirement (MPER), and empty body weight change (EBWC). The number of observations for these RFI components along lactation are different due to the high costs for collection. Therefore, there is a need to predict daily values of RFI components using appropriate statistical modeling.

Manafiazar et al (2013) developed a method for RFI computation based on a log-likelihood ratio test and Bayesian information criterion for component traits of RFI and multiple linear regression to obtain RFI values. The proposed method by Manafiazar et al (2013) for estimation of RFI in dairy cows was probably the most acceptable approach by researchers. For example,

this approach has been used by independent researchers for estimation of RFI in international collaborations from the United States, Canada, the Netherlands, and the United Kingdom (Hardie et al , 2017). They took advantage of pooled data from 4916 lactating cows to estimate phenotypic and genetic parameters of RFI. To date, this was the largest sample size that was used for estimation of RFI.

2.2 Materials and Methods

Experimental procedures were approved by the University of Alberta Animal Care and Use Committee for Livestock (AUP00000170). Dairy cows were handled in accordance with the guidelines described by the Canadian Council on Animal Care (CCAC 2020). The animal study was conducted at Dairy Research and Technology Center (DRTC).

2.2.1 Barn management

Mixed parity lactating Holstein cows were used for this study. Milk production records of 75 dairy cows from June 2017 to October 2018 were used in the analysis. Animals were kept in a tie-stall system at the Dairy Research and Technology Center (DRTC) at the University of Alberta. Ration at two levels of energy, high and medium, was fed to animals between the first days of lactation and day 240 in milk (DIM) as total mixed ration (TMR) based on their milk production level. Daily records of offered feed intake to each animal were taken at the morning and the weight of refused feed was weighed the morning of the next day.

All cows in the project were fed the same diet after calving. The dry off diet was usually offered to the dairy cows when they are close to parturition and contained approximately 20% concentrate on a dry matter basis. After parturition cows were gradually switched during the first

seven days to an early lactation diet with higher proportion of grain (up to 50% on DM basis) to meet the energy demands for high milk production cows. Then animals moved to mid ration as their milk production declines. Daily ration was offered as total mixed ration (TMR) for ad libitum intake to allow approximately 5% feed refusal throughout the experiment. Animals were fed once daily in the morning at 8:00 AM.

2.2.2 Blood sampling

Blood samples were taken at three time points: early, mid and late lactation (50, 150, 240 DIM) from all 75 cows. The blood samples were obtained from the coccygeal vein 6:00 A.M before feeding. All blood samples were collected in 10mL vacutainer tubes and were allowed to clot by keeping at room temperature until separation of serum. Clotted blood samples were centrifuged at 2090 × g at 4 °C for 20 minutes (Rotana 460 R centrifuge, Hettich Zentrifugan, Tuttlian, Germany) and the separated serum were aspirated and transferred to a sterile 2mL tube. Finally, all serum samples were stored at -80 °C until metabolomics analysis.

2.2.3 Data Editing and RFI component traits calculation

Three cows were removed from analysis because they did not fall within three standard deviations from the population mean on the test day. The remaining 72 cows had enough number of records and met the criteria of normal distribution to be analyzed for RFI estimation using random regression model (RRM).

RFI is equal to the difference between actual energy intake (AEI) and expected energy intake (EEI), the formula is as follow:

$$RFI = AEI – EEI$$

Residual feed intake is the actual energy intake that is accounted for energy requirements of the cow. Individual actual energy intake can be calculated with net energy density (ED), offered feed, refused feed and dry matter (DM) percentage. DMI was calculated as:

DMI (Kg/day) = [offered feed (Kg) – refused feed (Kg)] × DM %

AEI (Mcal/day) = DMI (Kg/day) × ED (Mcal/kg)

ED is the amount of energy per kilogram dry matter of ration (Mcal/ kg DM).

Expected energy intake (EEI) has several components including metabolic body weight (MBW), milk production energy requirement (MPER) and empty body weight change (EBWC). Here we explain how each of these components were calculated:

1) Metabolic body weight (MBW) was defined as (NRC, 2001):

$$\text{MBW (kg)} = BW^{0.75}$$

Where BW is body weight (kg) of the animal which was measured monthly at 29th each month in the morning 6 AM. The analysis of BW records between 5 to 240 day in milk (DIM) were not adjusted for fetus growth weight because based on the fetus growth energy requirement is negligible in this period of pregnancy (NRC, 2001).

2) Another component of expected energy intake (EEI) is milk production energy requirement (MPER) which can be calculated using the following formula:

MPER (Mcal/day) = [(0.0929 × fat %) + (0.0547 × Crude Protein %) + (0.0395 × Lactose %)] × Milk Yield.

As the above formula shows the total energy that is expressed in the form of milk is included in MPER. The amount of milk yield, concentration of fat, protein and lactose are also considered in this formula (Eastridge 2002, Linn 2003, Zanton 2016). Milk yield was measured daily in

morning and afternoon milking by an automatic milking machine. Milk samples of AM and PM milking was sent to a central milk testing laboratory (CMTL).

3) To avoid the confounding effect of water and gastrointestinal content with body weight (Bewley et al., 2008), body weight is needed to be adjusted for gut fill (GF) as following formula shows

EBW (kg) = BW – GF,

GF (kg) = DMI × (11-(7×MED/15)),

Where MED is metabolizable energy density (Mcal/Kg).

2.2.4 Random Regression Model (RRM) for RFI component traits

After calculation of daily AEI, monthly MBW, weekly MPER and monthly EBW which are all elements to calculate RFI, it is needed to predict MBW, MPER and EBW values and smooth AEI. For this purpose a random regression model (RRM) was used (Manafiazar et al. 2013). Multivariate mixed model was as follow:

$$y_{it} = F_{it} + \sum_{m=0}^{k1} \beta_m P_m(t) + \sum_{m=0}^{k2} \lambda_{im} P_m(t) + \varepsilon_{it}$$

Where y_{it} is ith animal record on time (t)

F_{it} is contemporary grouping of animal i on time (t) which is a fixed effect in the model

β_m is regression coefficient for fixed effect trend of order m

λ_{im} is random regression coefficient of animal's permanent environmental effect plus additive genetic effect.

$P_m(t)$ is the mth order of time (t)

K1 and K2 are the orders of fixed effects and random effects regression respectively

Contemporary grouping effects were combined month and year of measurement, ration type, temperature and humidity index (THI), lactation number and age at first calving (AFC) as covariate.

The assumptions of the models such as normality and homogeneity of variance, were checked with diagnostic plots. The best model was selected using log-likelihood ratio test (LRT) and Bayesian information criterion (BIC).

2.2.5 Multiple linear regression for computation of RFI

Prediction of RFI component traits were done with RRM and daily value for these items which are predicted metabolic body weight (PMB), predicted empty body weight (PEBW) and predicted milk production energy requirement (PMPR) and smoothed actual energy intake (AEI) are gained. Therefore, we have the required component traits for prediction of RFI. It should be noted that the merit of feed efficiency is calculated based on the whole experiment from day 3 to 240 of lactation.

For computation of RFI, empty body weight change (EBWC) was used as the difference between two consecutive days.

EBWC (t) = EBW (t) – EBW (t-1)

The model for computation of RFI is:

$$\sum_{i=3}^{240} \mathrm{AEI_i} = P_i + \beta_0 + \beta_1 \sum_{i=3}^{240} MBW_i + \beta_2 \sum_{i=3}^{240} MPER_i + \beta_3 \sum_{i=3}^{240} EBWC_i + \sum_{i=3}^{240} RFI_i \, .$$

Where $\sum_3^{240} AEI_i$ is the summation of actual energy intake for ith animal between day 5 and 240 in milk (DIM), $parity_i$is the parity of ith cows, $\sum_3^{240} MBW_i$ is the summation of metabolic body weight (MBW) for ith animal between day 3 and 240 in milk (DIM), $\sum_3^{240} EBWC_i$ is the

summation of empty body weight change (EBWC) for ith animal between day 3 and 240 in milk (DIM), $\sum_{3}^{240} MPER_i$ is the summation of milk production energy requirement (MPER) for ith animal between day 5 and 240 in milk (DIM) and $\sum_{3}^{240} RFI_i$ is the residual of them model which will give us the cumulative RFI for ith animal along the lactation DIM (3-240).

All statistical analysis for RFI computation were performed using the MIXED procedure of SAS (SAS Institute, Inc. 2003).

All animals were sorted based on the RFI values and animals with RFI value over half a standard deviation from the mean on positive (μ + 0.5 SD) and negative (μ - 0.5 SD) sides were selected as high and low RFI animals for metabolomics analysis.

RRM modeling with five fixed and five random effects model made 25 models. AIC, BIC and log likelihood values were used to compare models. The model with order of 2 for fixed effects and order of five for random effects had the best fit.

Multiple linear regression for calculation of cumulative RFI had R-squared of 0.71. Therefore 29% of variation of dry matter intake (DMI) was not explained by derived traits. The daily derived traits such as AEI, PMBW, PMPER, PEBW, PEBWC as it has biological importance, were plotted along days in milk. Residual feed intake was calculated as the average of 3-240 DIM for each cow. RFI had normal distribution and a range from -5.06 to +4.83 on average in 240 days of lactation. Positive and negative RFI values are symmetrically distributed around 0 as figure 2.1 shows.

RFI had an average of 0.03 with a normal distribution which means half of animals with positive RFI value at the right side of the plot (Figure 2.1) are feed inefficient in comparison to the average of the herd (72 cows). Phenotypic standard deviation of RFI was 2.13 for 72 cows.

As the figure 2.1 shows, the 72 RFI predicted cows were ranked and categorized into most efficient (low RFI, RFI < 0.5 SD from the mean) and least efficient (high RFI, RFI > 0.5 SD from the mean). Then, serum samples from high RFI (n=20) and low RFI (n=20) cows were sorted out at all sampling time points (50, 150 and 240 DIM) and analysed using NMR metabolomics approach.

Lactating cows in high and low RFI groups show different trends in the components of RFI prediction equation. Actual energy intake (AEI), MBW, EBWC, MPER were components of RFI calculation equation. The change in the RFI components were plotted during lactation for high and low RFI groups.

2.3 Results

2.3.1 Actual energy intake (AEI)

Blood sampling days which are 50, 150 and 240 days in milk (DIM) are indicated by the vertical lines (Figure 2.2). Blue and red lines are the average of actual energy intake of low and high RFI group cows, respectively.

The energy intake rose rapidly from day 3 to 50 of lactation, then its growth slows down but still with an overall positive slope increased till 150 DIM that is the second sampling time point. Late lactation had almost fixed trend with fluctuations till the last day of blood sampling. None of the days had significant different actual energy intake between high and low RFI groups.

2.3.2 Milk Production Energy Requirement (PMPER)

Milk production energy requirement as the MPER formula (mentioned in section 2.2.3) is a function of milk yield, fat, protein and lactose concentration. Predicted milk production energy

requirement, which is modeled based on weekly MPER observations, increased rapidly in early lactation and gets to the plateau around day 40 of lactation (Figure 2.3). Then PMPER increases steadily up to day 150 of lactation and with a small drop arrives to day 240 of lactation. The gap between high and low RFI groups is clear after day 50 of lactation. There were no significant differences between groups in PMPER along lactation.

2.3.3 Predicted Empty Body Weight (PEBW)

Predicted empty body weight shows a nonlinear trend of body weight along the lactation (Figure 2.4). As lactation starts body weight is decreasing and both RFI groups experience their minimum somewhere in the vicinity of day 50 of lactation. The trend has fluctuations and not a clear greater value from either RFI group. Lactating cows recover their body weight with considerable slope from day 50 to 150. Then this body weight gain slows down, however cows are still increasing their body reserves up to the end of the experiment. The difference of PEBW between groups was not significant in any days of the experiment.

2.3.4 Predicted metabolic body weight (PMBW)

Predicted metabolic body weight trend shows a small reduction from start of lactation to day 50 (Figure 2.5). Then it increased from day 50 to 150 with a gap between high and low RFI groups. The gap in metabolic body weight is constant between groups in all periods of the experiment. After day 150, PMBW grows with lesser slope in comparison to the mid-lactation stage (day 50-150).

2.3.5 Predicted empty body weight change (PEBWC)

PEBWC shows a fluctuating trend at the beginning before 50 DIM. It increases up to 125 DIM and then is reduced up to 200 DIM. The last 40 days of experiment represents another fluctuation in predicted metabolic weight. There is almost no clear gap between groups along the lactation for this component trait.

2.3.6 Weekly average RFI value for high and low RFI groups

Weekly average of RFI for both groups is demonstrated in figure 2.7. It shows a constant fluctuation for both groups. The gap is clear between groups and was significant ($P < 0.05$) in all time intervals. The 50, 150 and 240 DIM are shown as vertical lines.

2.4 Discussion

Based on the definition of RFI, there is a difference between our expectation from cow energy requirement and her actual energy intake. This has been noticed since a long time that some of animals eat less and some pretty close to the expectation of energy requirement and some have appetite to consume more (Koch et. al 1963). It has caught the attention of researchers to measure this variability to consider it as a measure of feed efficiency because part of that phenotypic variation can be heritable and be useable in breeding programs. The reason behind this variability has been shown to originate from differences in metabolism in beef and dairy cows. Selection for high and low RFI bulls in beef cattle and the study on next generation provided evidence that protein turnover, tissue metabolism, body composition, and heat increment of fermentation explain most of variation of RFI. This study brought the attention to

role of metabolism in RFI regulation and led researchers to investigate feed efficiency in dairy cows with a metabolomics approach (Richardson and Herd 2004, Wang and Kadarmideen 2019). In this study, we considered divergent RFI groups of lactating cows for daily changes of RFI and its component traits such as MPER, MBW, EBWC and AEI. This can show the differences in energy metabolism between high and low RFI cows along lactation. Even though other researchers used random regression models for estimation of RFI in dairy, they did not show the trajectory of RFI and its derived traits for divergent RFI groups (Manafiazar 2013, Hardie et al. 2017, Li et al. 2020)

Normality of the RFI trait is one of the important features of this trait. Not only does this show how the feed efficiency merit is randomly distributed between animals, but it also facilitates downstream phenotypic and genetic analysis. Ranked animals by RFI values (Figure 2.1) showed a distribution which is in accordance with plots of mean of RFI per sire group and average daily residual feed intake from lactating cows (Davis et al. 2014, Manafiazar et al. 2013). Hassoun et. al (2019) also ranked dairy ewes by RFI values and showed the same distribution. The histogram of (Figure 2.1) is a normal bell shape distribution because most of animals have RFI close to zero. The most feed efficient lactating cow (Figure 2.1) with RFI of -5.06 Mcal/day at the left side shows that this cow consumes 5.06 Mcal less energy per day than average of 72 cows from day 3 to 240 of lactation. Part of this merit in feed efficiency comes from genetic basis and animals with negative RFI values can be selected to be parents of next generation. The groups of cows that are selected for metabolomics study are shown in (Figure 2.1).

Actual energy intake (AEI) of lactating cows in low and high RFI groups showed a clear and significant ($P < 0.05$) separation between the groups (Figure 2.2). This figure shows a sharp increase of actual energy intake in the first 100 days and the steady mild trend after that until the

end of experiment. Actual energy intake traditionally was a trait of interest for its importance in energy balance and health traits of dairy cows during the transition period (Collard et al. 2000) However in recent years, it has become a trait of interest for computation of RFI in beef and dairy cattle (Berry and Crowley 2013, Pryce et al. 2015). Direct measurement of individual feed intake of dairy cows along lactation is expensive and has been one of the largest challenges in improvement of feed efficiency (Connor, 2014). Recent studies have shown a similar trend in AEI and DMI in dairy cows and our result agrees well with these reports (Manafiazar et al. 2013; Li et al. 2020).

Predicted milk production energy requirement (PMPER) of high and low RFI cows grew rapidly in both groups (Figure 2.3) from first days of lactation with a gap. The time point close to maximum energy requirement (50 DIM) is associated with higher metabolism load for increased level of milk production. PMPER stayed nearly constant in mid-lactation and then it falls as the cow gets closer to the end of lactation. High and low RFI cows that have different energy efficiency classes may have differences in metabolism at these days that metabolic pressure for production is higher (Figure 2.3). The blood samples that were taken at this day are analyzed in chapter 3.

Empty body weight (EBW) is an adjusted body weight for gut fill (VandeHaar and St-Pierre 2006; Yu et al. 2003). High and low RFI groups had very close PEBW up to the 50 DIM where the cows experience the minimum weight along the 240 days of lactation. Then body weight recovered gradually with steady increase after 150 DIM. This (Figure 2.4) shows how divergent groups of RFI cows handle the metabolic pressure of catabolism in early lactation and anabolism in mid- and late-lactation stage. The difference between the two RFI groups is not significant and the trends do not show a clear gap between groups except from the day 130 to190. In this time

interval, low RFI group cows seem to be steadily gaining more body weight than the high RFI group. The day 150 of sampling for metabolomics study is almost in the middle of this period which has potential to show biological differences between high and low RFI groups. Body weight gain continued up to day 240 of lactation with considerably less rate of changes in comparison to the range 50-150 DIM.

Metabolic body weight (MBW) is interpreted as the weight of body tissues that are active in metabolism and has been used in dairy industry for decades for estimation of maintenance energy requirements. (Gaines, 1946). From day 3 to 50 of lactation, metabolic body weight does not show lots of changes (Figure 2.5), however, a mild fall and rise is evident. MBW increased considerably after day 50 of lactation, and then it reaches its plateau by day 240.

MBW was numerically different between RFI groups along the lactation, however this difference was not significant ($P > 0.05$). The differences are not statistically significant, but the trend is important to be observed for tracking the metabolism trend along lactation in high and low RFI groups.

PEBWC is probably the most interesting result from this visualization. PEBWC is plotted (figure 2.6) for both groups and shows they lose around 0.5 kg body weight every day in the first 50 days of lactation. After day 50 both RFI groups pass the zero line (Figure 2.6) with a one-day gap. High RFI cows stop losing weight at 71 DIM while low RFI cows experience it at 72 DIM.

Weekly RFI values are demonstrated in Figure 2.7. This plot has a horizontal line at zero to show when RFI groups get negative or positive. The difference between high and low RFI groups have been statistically significant all along lactation. In overall, both groups show negative RFI values up to 40 DIM due to little dry matter intake in comparison to animal needs for milk production.

Limited capacity of rumen after calving and its gradual expansion to take more feed per day is the underlying reason for negative RFI for both groups in the first 50 days of lactation. Mid-lactation changes show more fluctuation around zero, however the gap between groups is still clear and wide. From day 150 to day 240, even low RFI cows experience positive values as they need to store energy for pregnancy requirements and the next lactation. To the best of our knowledge Li et al. (2020) has been the only study that showed the trend of daily RFI values using the data from USDA. The same fluctuations described above took place in the RFI results for the dairy cows reported by Li et al. (2020).

2.5 Conclusion

Overall, this chapter showed how RFI was estimated in dairy cows using body weight, milk production, fat, protein and lactose concentrations. Animals were sorted by RFI values and were grouped by high and low RFI groups. Changes of AEI, PMPER, MBW, EBWC and RFI were considered along lactation. Differences of high and low RFI groups were visualized for RFI and its derived traits. The differences between component traits were not significant. Actual energy intake (AEI) and RFI was significantly different between RFI groups as was shown before in lactating Holsteins (Elolimy, et al. 2018). Previous studies considered across experiment time interval for the differences of high and low RFI groups, but we used the same method for daily differences between groups in a longer period (DIM 0-240). The results were plotted for the period to see how they change by time. This information may be useful in biological interpretation of energy expenditure in divergent RFI groups of lactating cows and also shows the consistency of differences visually.

At critical time points of metabolism along lactation such as peak of lactation and minimum of body weight, there could be differences in metabolism. Previous studies on RFI were more focused on overall feed efficiency of lactating cows with a single biological sample for tacking differences in metabolism. This study could visualize the trend of changes of RFI along lactation. It can be useful for future studies on RFI groups that are interested to study biological differences with focus on metabolism status changes along lactation. Further investigation of biological differences between high and low RFI groups can shed light on some of the underlying reasons of RFI variation in dairy cows.

2.6 References

Arthur, J. P. F., & Herd, R. M. (2008). Residual feed intake in beef cattle. *Revista Brasileira de Zootecnia*, *37*(SPECIALISSUE), 269–279. https://doi.org/10.1590/S1516-35982008001300031

Basarab, J. A., Beauchemin, K. A., Baron, V. S., Ominski, K. H., Guan, L. L., Miller, S. P., & Crowley, J. J. (2013). Reducing GHG emissions through genetic improvement for feed efficiency: effects on economically important traits and enteric methane production. *Animal : An International Journal of Animal Bioscience*. https://doi.org/10.1017/S1751731113000888

Berry, D. P., & Crowley, J. J. (2012). Residual intake and body weight gain: A new measure of efficiency in growing cattle. *Journal of Animal Science*, *90*(1), 109–115. https://doi.org/10.2527/jas.2011-4245

Berry, D. P., & Crowley, J. J. (2013). Cell biology symposium: Genetics of feed efficiency in dairy and beef cattle. *Journal of Animal Science*, *91*, 1594–1613.

https://doi.org/10.2527/jas2012-5862

CCAC. (2020). Retrieved July 15, 2020, from https://www.ccac.ca/

Collard, B. L., Boettcher, P. J., Dekkers, J. C. M., Petitclerc, D., & Schaeffer, L. R. (2000). Relationships between energy balance and health traits of dairy cattle in early lactation. *Journal of Dairy Science*, *83*(11), 2683–2690. https://doi.org/10.3168/jds.S0022-0302(00)75162-9

Connor, E. E. (2014). Invited review: Improving feed efficiency in dairy production: challenges and possibilities. https://doi.org/10.1017/S1751731114002997

Connor, E. E., Hutchison, J. L., Norman, H. D., Olson, K. M., Van Tassell, C. P., Leith, J. M., & Baldwin VI, R. L. (2013). Use of residual feed intake in Holsteins during early lactation shows potential to improve feed efficiency through genetic selection. *Journal of Animal Science*, *91*(8), 3978–3988. https://doi.org/10.2527/jas.2012-5977

Crews, D. H. D. (2005). Genetics of efficient feed utilization and national cattle evaluation: a review. *Genetics and Molecular Research : GMR*, *4*(2), 152–165. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16110437

Davis, S. R., Macdonald, K. A., Waghorn, G. C., & Spelman, R. J. (2014). Residual feed intake of lactating Holstein-Friesian cows predicted from high-density genotypes and phenotyping of growing heifers. *Journal of Dairy Science*, *97*(3), 1436–1445. https://doi.org/10.3168/jds.2013-7205

Eastridge, M. L. (2002). Energy in the New Dairy NRC (p. 11). Feedstuffs. Retrieved from https://dairy.osu.edu/sites/dairy/files/imce/PDF/Feed_PDF/Energy New Dairy NRC.pdf

Elolimy, A. A., Arroyo, J. M., Batistel, F., Iakiviak, M. A., & Loor, J. J. (2018). Association of residual feed intake with abundance of ruminal bacteria and biopolymer hydrolyzing

enzyme activities during the peripartal period and early lactation in Holstein dairy cows, 1–10. https://doi.org/10.1186/s40104-018-0258-9

Gaines, W. L. (1946). Live Weight Versus Metabolic Body Size in Dairy Cows and Goats. *Journal of Dairy Science*, *29*(5), 259–272. https://doi.org/10.3168/jds.S0022-0302(46)92477-0

Hardie, L. C., VandeHaar, M. J., Tempelman, R. J., Weigel, K. A., Armentano, L. E., Wiggans, G. R., … Spurlock, D. M. (2017). The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. *Journal of Dairy Science*, *100*(11), 9061–9075. https://doi.org/10.3168/jds.2017-12604

Hassoun, P. (2019). Residual feed intake in dairy ewes : an evidence of intraflock variability. *BioRxiv*, (August), 723809. https://doi.org/10.1101/723809

Koch, R. M., Chambers, D., & Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *Journal of Animal Science*, *22*(2), 486–494.

Korver, S., van Eekelen, E. A. M., Vos, H., Nieuwhof, G. J., & van Arendonk, J. A. M. (1991). Genetic parameters for feed intake and feed efficiency in growing dairy heifers. *Livestock Production Science*, *29*(1), 49–59. https://doi.org/10.1016/0301-6226(91)90119-B

Li, B., VanRaden, P. M., Guduk, E., O'Connell, J. R., Null, D. J., Connor, E. E., … Cole, J. B. (2020). Genomic prediction of residual feed intake in US Holstein dairy cattle. *Journal of Dairy Science*, *103*(3), 2477–2486. https://doi.org/10.3168/jds.2019-17332

Linn, J. (2003). Energy in the 2001 Dairy NRC : Understanding the System, (Figure 1), 102–110.

Manafiazar, G., Goonewardene, L., Miglior, F., Crews, D. H., Basarab, J. A., Okine, E., & Wang, Z. (2015a). Genetic and phenotypic correlations among feed efficiency, production and selected conformation traits in dairy cows. *Animal*, *10*(3).

https://doi.org/10.1017/S1751731115002281

Manafiazar, G., Goonewardene, L., Miglior, F., Crews, D. H., Basarab, J. A., Okine, E., & Wang, Z. (2015b). Genetic and phenotypic correlations among feed efficiency, production and selected conformation traits in dairy cows. *Animal*, *10*(3), 381–389. https://doi.org/10.1017/S1751731115002281

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., Basarab, J., Li, P., & Wang, Z. (2013). Prediction of residual feed intake for first-lactation dairy cows using orthogonal polynomial random regression. *Journal of Dairy Science*, *96*(12), 7991–8001. https://doi.org/10.3168/jds.2013-6560

Manafiazar, G. (2013). Investigation of Residual Feed Intake in Dairy. (Publication No. NS27639.) [PhD thesis, University of Alabama]. ProQuest Dissertations & ThesesGlobal.

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., & Wang, Z. (2012). Factors That Affect Energy Efficiency and Indirect Response of Selection for Efficiency on Related Traits. https://doi.org/http://dx.doi.org/10.5772/50776

McDonald, P., Edwards, R. A., Greenhalgh, J. F. D., & Morgan, C. A. (2002). Animal Nutrition. (7th Editio). Pearson Canada.

Pryce, J. E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W. J., Coffey, M. P., Hayes, B. J., & Goddard, M. E. (2015). Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. *Journal of Dairy Science*, *98*(10), 7340–7350. https://doi.org/10.3168/jds.2015-9621

Pryce, J. E., Wales, W. J., De Haas, Y., Veerkamp, R. F., & Hayes, B. J. (2014). Genomic selection for feed efficiency in dairy cattle. *Animal*, *8*(1), 1–10. https://doi.org/10.1017/S1751731113001687

Richardson, E. C., & Herd, R. M. B. (2004). Biological basis for variation in residual feed intake in beef cattle . 2 . Synthesis of results following divergent selection Cooperative Research Centre for Cattle and Beef Quality . *Australian Journal of Experimental Agriculture*, *44*, 431–440. https://doi.org/10.1071/EA02221

Vandehaar, M. J. (1998). Efficiency of Nutrient Use and Relationship to Profitability on Dairy Farms. *Journal of Dairy Science*, *81*(1), 272–282. https://doi.org/10.3168/jds.S0022-0302(98)75576-6

VandeHaar, M. J., & St-Pierre, N. (2006). Major Advances in Nutrition: Relevance to the Sustainability of the Dairy Industry. *Journal of Dairy Science*, *89*(4), 1280–1291. https://doi.org/10.3168/jds.S0022-0302(06)72196-8

Veerkamp, R. F., & Emmans, G. C. (1995). Sources of genetic variation in energetic efficiency of dairy cows. *Livestock Production Science*, *44*(2), 87–97. https://doi.org/10.1016/0301-6226(95)00065-0

Wang, X., & Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites*, *9*(7), 1–16. https://doi.org/10.3390/metabo9070151

Yu, P., Meier, J. A., Christensen, D. A., Rossnagel, B. G., & McKinnon, J. J. (2003). Using the NRC-2001 model and the DVE/OEB system to evaluate nutritive values of Harrington (malting-type) and Valier (feed-type) barley for ruminants. *Animal Feed Science and Technology*, *107*(1–4), 45–60. https://doi.org/10.1016/S0377-8401(03)00062-2

Zanton, G. I. (2016). Analysis of production responses to changing crude protein levels in lactating dairy cow diets when evaluated in continuous or change-over experimental designs. *Journal of Dairy Science*,*99*(6),4398–4410. https://doi.org/10.3168/jds.2015-10438

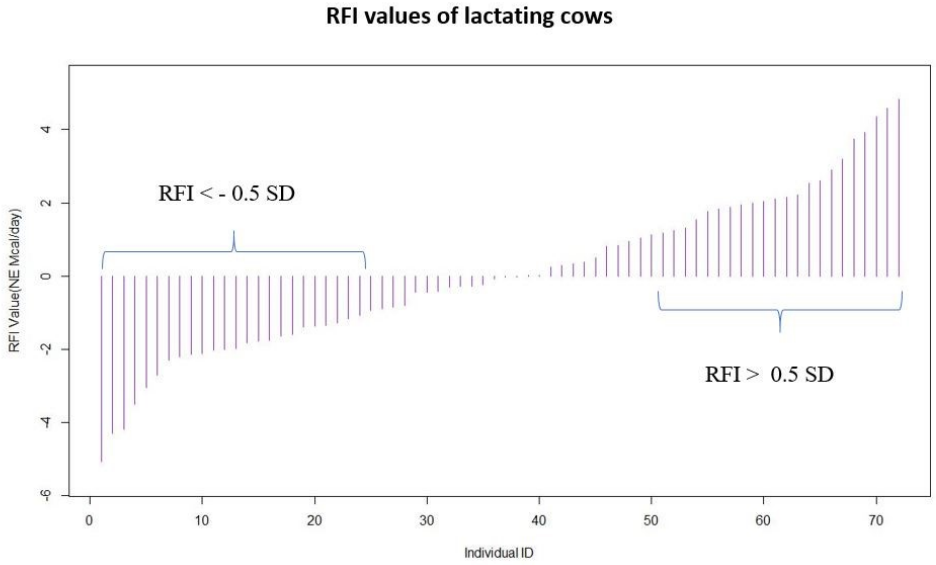

Figure 2.1 Lactating dairy cows were sorted by RFI values and the extreme individuals (< -0.5 SD and > +0.5 SD) were selected for metabolomics analysis.

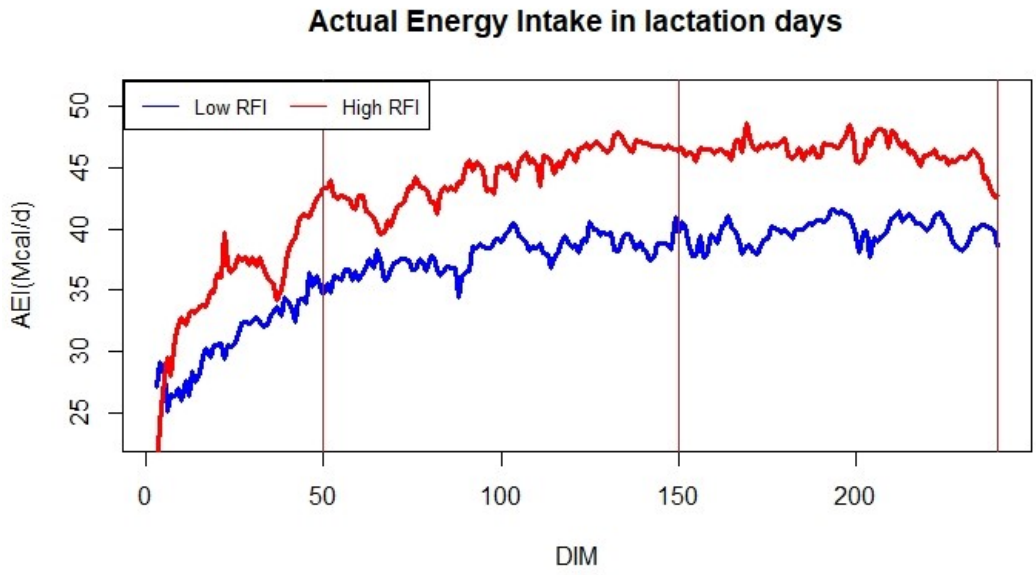

Figure 2.2 Actual Energy Intake (AEI) for both high and low RFI groups along lactation. Blood sampling days (50, 150, 240 DIM) are demonstrated as vertical lines.

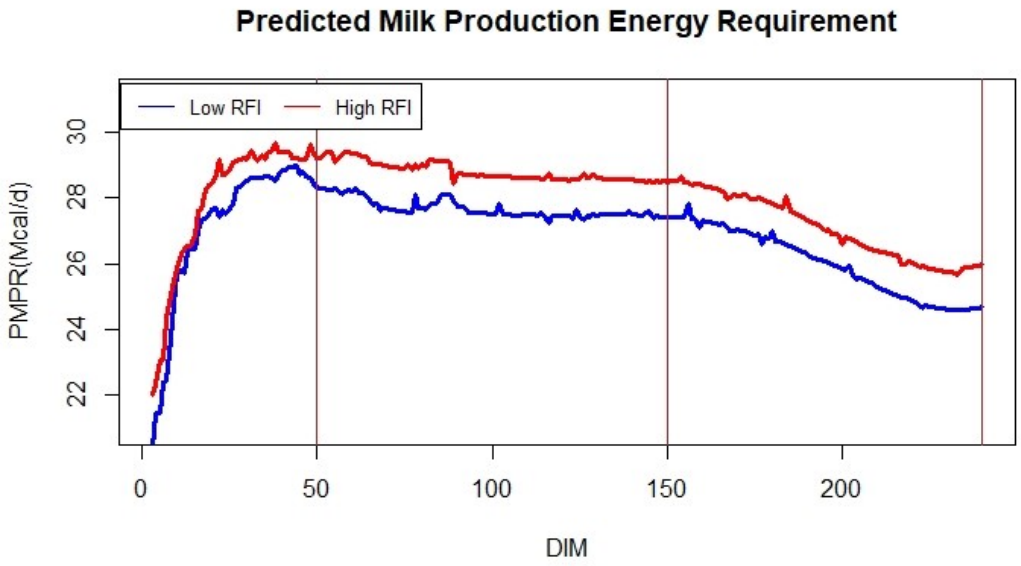

Figure 2.3 Predicted Milk Production Energy Requirement (PMPR) for both high and low RFI groups along lactation. Blood sampling days (50, 150, 240 DIM) are demonstrated as vertical lines.

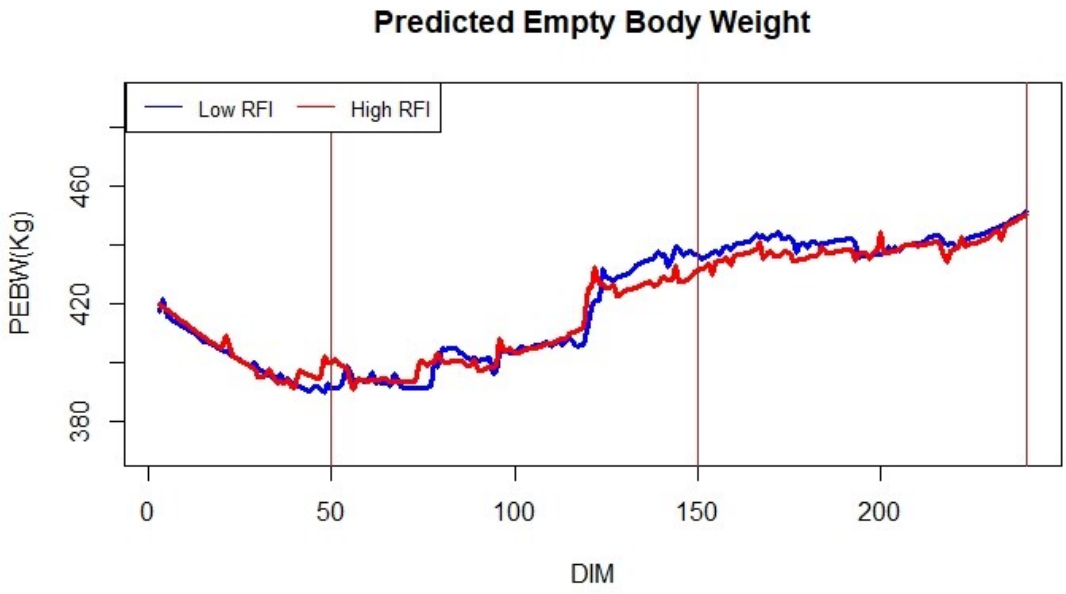

Figure 2.4 Predicted Empty Body Weight (PEBW) for both high and low RFI groups along lactation. Blood sampling days (50, 150, 240 DIM) are indicated as vertical lines.

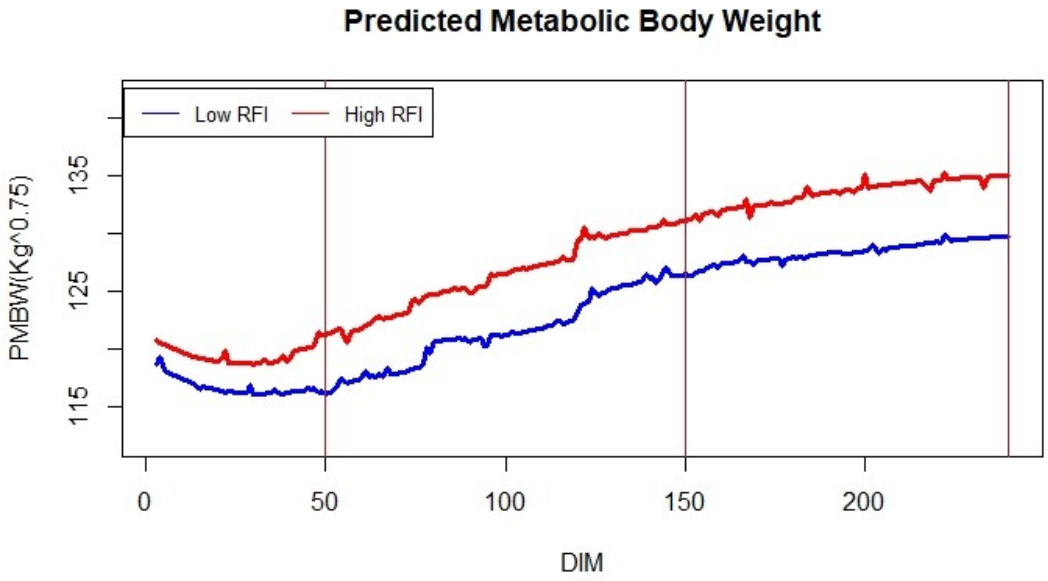

Figure 2.5 Predicted Metabolic Body Weight (PMBW) for both high and low RFI groups along lactation. Blood sampling days (50,150,240 DIM) are demonstrated as vertical lines.

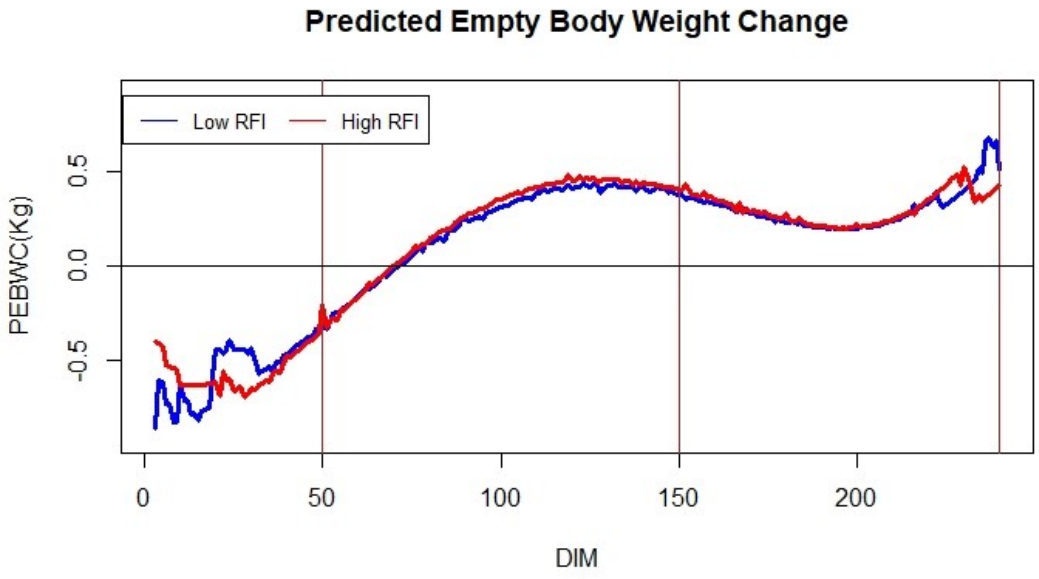

Figure 2.6 Predicted Empty Body Weight Change (PEBWC) for both high and low RFI groups along lactation. Blood sampling days (50,150,240 DIM) are demonstrated as vertical lines.

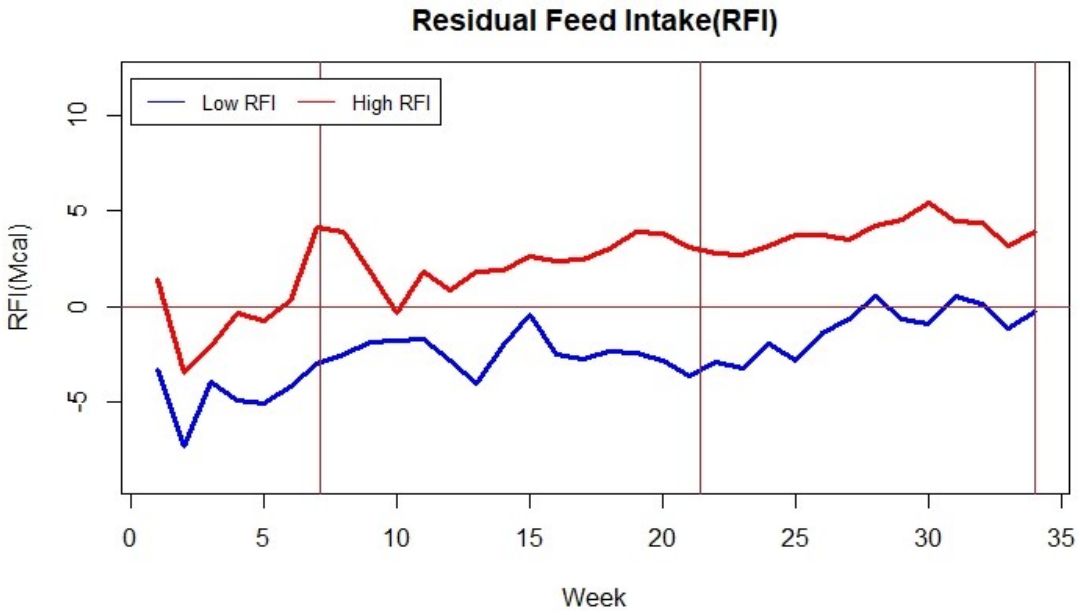

Figure 2.7 Residual Feed Intake for both high and low RFI groups along lactation. Blood sampling days (50,150,240 DIM) are demonstrated as vertical lines

Chapter 3. Metabolomics approach to uncover differences between high and low RFI dairy cows

3.1 introduction

Physiological mechanisms have significant contribution to the variation of RFI in beef and dairy cattle. The association of RFI with metabolite profile has been earlier studied in beef cattle than dairy cows. The experience of researchers from beef have shown that no single mechanism is primarily responsible for the variation in this trait (Herd & Arthur, 2009). Some of the pathways and physiological functions that contribute to RFI have been discovered in recent years. They are useful to increase the knowledge about the factors that affect efficiency of feed utilization in animals. Furthermore, tracking biological pathways may lead us to find genes that play significant role in regulation of RFI in animals (Karisa, et al. 2013).

One of the experiments that could show biological differences between high and low RFI animals was done using divergent selection for RFI in Angus beef cattle (Arthur et al. 2001). This study took a long time to measure RFI and related traits in progeny of two RFI groups but had interesting results. Arthur et al. (2001) demonstrated metabolic processes such as heat production from metabolic processes, body composition and physical activity altogether explains 73% of variation of RFI (Herd & Arthur, 2009).

These findings which had been gained by actual selection on animals brought the attention of researchers to experiments that could show biological differences between high and low RFI cattle groups. Therefore, multi-omics studies using different types of samples such as blood and liver with metabolomics and proteomics approaches were used to expand our knowledge of physiology of feed efficiency in beef cattle. Such studies also attempted to map metabolites and proteins to genes through biological pathways (Karisa et al. 2014, Fonseca et al. 2019). The goal

of such studies was to detect potential biomarkers that can be used to classify high and low RFI animals. The other outcome is that they also show how gene regulation is different in cell processes by biological pathways analysis. Moreover, interesting efforts of researchers has been done to develop models with molecular data that can explain variation of RFI and predict its values for individuals in the herd (Karisa et al., 2014).

Residual feed intake is a complex trait with moderate to low heritability that can be better understood with omics studies than only GWAS approach. There are also potentials that complex traits could be explained with statistical models that use concentration of biological molecules as predictors (Kadarmideen 2008). Given that these molecular data (metabolites or proteins) are heritable and relatively inexpensive to measure compared to directly calculating residual feed intake, there would be chances for molecular data to be used for genetic selection to improve complex traits in dairy cattle.

Even though there are basic differences in energy partitioning between beef and dairy cows, similar metabolic processes are likely to have significant influence on variation of RFI. Yao et al. (2013) showed a genetic basis for this similarity. They mapped 38 SNPs to RFI quantitative loci regions and four of them were previously detected by GWAS studies in beef cattle. Wang and Kadarmideen (2019) studied RFI in two dairy cattle breeds and found important biological networks associated with feed efficiency. However, the latter study used a small sample size of 5 animals per group of RFI which resulted in limited findings. The other aspect that is not known yet is the biological differences in different stages of lactation between two RFI groups of lactating cows. Thus, our project has been defined to investigate differences between high and low RFI Holstein cows with sample size of 20 per group of animals in 3 stages of lactation. Larger sample size will most likely show more differences because of higher degree of freedom

that increases the power of the test. Moreover, samples are taken at day 50, 150 and 240 DIM representing the early, mid and late stage of lactation.

One of the differences between metabolomics study and other omics is the nature of sample type. Metabolomics needs body fluids like urine, rumen fluid, blood etc. However, although metabolomics has been developed in tissues like liver, brain and other organs in different species, originally the method was used for body biofluids (Wang et al. 2019). The question of which fluid is better for a metabolomic study depends on which fluid is more involved with the phenotype of interest. RFI is a measure of feed efficiency and different biochemical reactions play roles in its expression (Manafiazar et al., 2012). From digestion and absorption of the feed to its conversion to energy for the body or the application of energy in basal metabolism and production, many tissues are involved in efficient use of energy (Nkrumah, 2006). All organs of digestive system such as rumen, omasum, abomasum and intestine work together to extract feed energy for animal maintenance and production requirements. Then part of energy is used for maintenance requirements of cows. This involves any organ that need energy to function which means all of the body. Part of the energy is saved as body reserves, which will be used at the time of need in her peak of lactation or when she does not have access to enough feed (Moe, 1981).

The energy in biological environment is hidden in chemical bonds of molecules. The small molecules can leave one organ to the blood stream and release energy in another organ. Small molecules in the body play a vital role in transportation of nutrients and intermediate end products which primarily are carried by blood. Every tissue and any organ are bathed in blood and cell by-products are carried by blood to kidney and consequently bladder for excretion.

Blood as a fluid is mentioned as a liquid highway for molecules that are products of secretion or excretion molecules in response to physiological needs (Serkova et al., 2011).

Furthermore, metabolomic data from blood have already shown the ability to predict RFI values in beef cattle. Karisa et al. (2014) reported 48% and 99.5% accuracy of prediction for models that had three and twelve metabolites respectively as predictors of RFI. The application of blood metabolites in early diagnosis of diseases in transition dairy cows is another example of metabolites potential functionality in future livestock industry to prevent severe stages of diseases in dairy cows. This may reduce health-related economic cost of production such as mastitis, metritis, ketosis etc. given the metabolomic diagnosis approach can be done with reasonable cost. In a recent study on transition dairy cows done by Hailemariam et al. (2014), predictability of periparturient diseases four weeks before appearance of clinical symptoms was reported using blood metabolites. They developed a model that could predict periparturient diseases using only three blood metabolites with sensitivity and specificity of 87% and 85% (Hailemariam et al., 2014).

These facts have made blood the most interesting fluid for metabolomics study in livestock. Most of studies that are done with metabolomics approach in beef and dairy have used serum and plasma, respectively in 21 and 14 papers. These research articles have been used as references in the Livestock Metabolome Data Base (LMDB) (Goldansaz et al., 2017). In our study, we analysed serum metabolites since it reflects the metabolism in wide variety of tissues that are important for feed efficiency. Furthermore, literature (Karisa et al. 2014) in beef studies supports the usefulness of blood fluid for investigation of RFI.

A look at metabolomics studies that are done to investigate RFI shows that there is no specific preferred fluid of blood (serum or plasma) that provides significantly different results, however,

there are limited studies for this specific trait on beef/dairy cow. Nevertheless, it is possible to consider human studies that have already used NMR technology to evaluate differences between plasma and serum sensitivity and stability regardless of the case of study. Yu et al. (2011) considered concentrations of 122 metabolites with over limit of detection, each serum and plasma sample were analyzed for metabolite profiling twice. The reason to perform metabolomics analysis using NMR twice was to compare stability and sensitivity of blood plasma and serum. Mean correlation of all considered metabolites between the replicates were reported as 0.80 and 0.83 for serum and plasma which shows relatively better stability of plasma samples. Their result also indicated that 104 metabolites of serum have significantly more concentrations than plasma. Yu et al. (2011) concluded that the higher concentration of metabolites in serum samples makes biomarker discovery studies more sensitive for detection of differences between experimental groups.

Different energy transformations take place during lactation. Anabolism and catabolism of body tissues to support production, and digestion and absorption of feed are the main metabolic processes that dairy cows undergo during lactation (Martin & Sauvant, 2007). Numerous researchers believe variation in RFI is the result of variation in metabolism of animals by providing pathways that are regulated significantly differently between high and low RFI groups (Salleh et al. 2017, Khansefid et al. 2017).

In this project, I took blood samples from lactating dairy cows in early, mid, and late lactation stages to investigate potential differences in metabolism between high and low RFI groups. The next section of introduction will explain the typical production rhythm of lactating dairy cows in early, mid, and late lactation. The three important stages of lactation that are not yet investigated if there is any significant difference in metabolite concentration between high and low RFI

animals. The day 50, 150, and 240 of DIM are selected as a day roughly in the middle of these stages of lactation.

3.1.1 Metabolism of lactating cows in stages of lactation

Lactating dairy cows start lactation with a rise of milk yield that continues up to the peak of lactation that happens around week 6-8 of lactation. Then, milk production gradually reduces to the end of lactation. In this first three months of lactation, called “early lactation”, dairy cows are unable to eat as much as they need to provide the energy required to support milk production. This is due in part to changes in gut capacity as the body cavity has been extensively occupied by the embryo for the final months before calving. The rumen also reduces during pregnancy specifically in last two month. That means rumen digestion capability and the amount of feed intake has reduced because of physical and physiological changes in pregnancy. After calving, the rumen starts expanding gradually to reach to its normal (non-pregnant) digestion ability. In early lactation, the rumen is not capable of taking as much feed as cow needs for peak of milk production (Neves et al., 2018). Therefore, the cow breaks down body tissues to provide fat and protein for the energy needed for high milk production. Energy turnover efficiency in this stage is still under study (Wang & Kadarmideen, 2019).

The next stage is termed mid-lactation and feed intake meets the energy requirement. The animal then starts storing energy in body tissues while milk production is being reduced. At this stage, the cow is less challenged metabolically in comparison to early lactation and body reserves are increasing. The peak of dry matter intake (DMI) happens between months 3 to 7 of lactation (French, 2017). Clearly, this period is the time when the cow recovers her energy sources partly

and with considerable body weight changes. Cow body weight sill increases in next stage of lactation to store energy for next milk cycle.

Late lactation is defined as after month 7 of lactation when the animal is still gaining body weight with less rate of change relative to the previous, mid-lactation, stage (O'Sullivan et al., 2020). Milk production is substantially decreased, and the concentration of milk elements like fat and protein are increased. This stage continues up to month 10 of lactation.

The last stage is termed the dry-off period and is after calving from month 10 to 12 and cows are not milked at this time. Weight gain is very noticeable during this period. A significant part of her body weight gain is due to supporting embryo requirements as it is in its main developmental stage of the embryo. This is a critical part of dairy herd management. Milk fever is a common consequence of poor management and nutrition in this stage which may lead to reduced feed intake, negative energy balance and consequently ketosis (Mulligan et al. 2006). Depending on body condition of cows and her health status, appropriate ration energy and ingredients needs to be considered in this period.

In this study, we aimed to observe physiological and metabolic differences between high and low RFI lactating Holsteins. It was our goal to find metabolites that can predict whether the animal is high or low RFI with good accuracy. As one of the purposes of our study was to detect potential metabolite biomarkers associated with RFI, serum was selected for metabolomics analysis in this study. We assumed differences in RFI of dairy cows have roots in differences in their metabolism and RFI class of animals (high/low) and the RFI values are predictable using metabolite profile of animals. To test these hypotheses, we used multiple linear regression to see if metabolite concentrations of RFI classes are significantly different in presence of other effects such as month-year of sampling, parity, etc. The prediction ability of metabolite profile

metabolites for classification of cows by RFI classes was tested using PLS-DA model and demonstrated with ROC plot. Multiple linear regression with stepwise method was used to assess the linear relationship of metabolite concentrations and RFI values of individuals.

3.2 Materials and methods

3.2.1 Sample preparation and NMR spectroscopy

Serum samples of high and low RFI dairy cows collected at 50, 150 and 240 DIM were analysed using NMR spectroscopy at The Metabolomics Innovation Centre (TMIC). A total of 120 serum samples (20 serum samples per group at three time points during lactation) were analysed. Serum samples contain a significant concentration of large molecular weight proteins and lipoproteins which affects the identification of the small molecular weight metabolites by NMR spectroscopy. A deproteinization step, involving ultra-filtration as previously described by Psychogios et al. (2011) was therefore introduced in the protocol to remove serum proteins. Prior to filtration, 3 KDa cut-off centrifugal filter units (Amicon Microcon YM-3), were rinsed five times each with 0.5 mL of H_2O and centrifuged (10,000 rpm for 10 minutes) to remove residual glycerol bound to the filter membranes. Aliquots of each serum sample were then transferred into the centrifuge filter devices and spun (10,000 rpm for 20 minutes) to remove macromolecules (primarily protein and lipoproteins) from the sample. The filtrates were checked visually for any evidence that the membrane was compromised and for these samples the filtration process was repeated with a different filter and the filtrate inspected again. The de-proteinized serum samples (280 μL) were then transferred to a 1.5 mL Eppendorf tube, to an additional 70 μL of the standard NMR buffer solution (54% D_2O:46% 1.75 mM KH_2PO_4 pH 7.0 v/v containing 5.84 mM DSS (2,2-dimethyl-2-silcepentane-5-sulphonate)).

The serum sample (350 µL) was then transferred to a 3mm SampleJet NMR tube for subsequent spectral analysis. All 1H-NMR spectra were collected on a 700 MHz Avance III (Bruker) spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) cryoprobe. 1H-NMR spectra were acquired at 25°C using the first transient of the NOESY pre-saturation pulse sequence (noesy1dpr), chosen for its high degree of quantitative accuracy (Saude, Slupsky, & Sykes, 2006). All FID's (free induction decays) were zero-filled to 250 K data points. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm) and quantification of all 1H-NMR spectra were processed and analyzed using an in-house version of the MAGMET automated analysis software package using a custom metabolite library. MAGMET allows for qualitative and quantitative analysis of an NMR spectrum by automatically fitting spectral signatures from an internal database to the spectrum. Typically, all visible peaks were assigned. Most of the visible peaks are annotated with a compound name. It has been previously shown that this fitting procedure provides absolute concentration accuracy of 90% or better (Ravanbakhsh et al., 2015).

3.2.2 Univariate metabolite analysis

The NMR panel for serum samples included 33 metabolites above limit of quantification (LOQ) for analysis. The LOQ is the lowest concentration that the metabolite can be not only reliably detected but also it does not have the imprecision and bias (Armbruster & Pry, 2008). Therefore, we used LOQ as the lowest acceptable metabolite concentration. The multiple linear regression assumptions of normality, homogenous variance, linear relationship of variables with response and other criteria were checked according to statistical literature standards and guidelines (Akinwande, et al. 2015). One of the important assumption of linear regression in normality of

data. Therefore, the distribution of metabolites concentration was checked before fitting the model.

Metabolite concentrations were checked to be normal before stepwise multiple linear regression analysis. In case that the distribution of metabolite concentration was not normal, the logarithm on base 10 transformation was applied to meet the assumptions of linear regression model. This transformation is usual in data processing before model development with metabolomics data (Thévenot, et al. 2015) to meet the assumption of the model. All data management and model development was done using R (R-Core-Team, 2019).

Metabolite concentration was taken as dependent variable to assess the effect of environmental factors and grouping effect of animals. The model to assess the effects was:

$$\boldsymbol{Metabolite_i} = \boldsymbol{MY}_i + \boldsymbol{AFC}_i + \boldsymbol{Parity}_i + \boldsymbol{Health}_i + \boldsymbol{Ration}_i + \boldsymbol{RFI_i} + \boldsymbol{e_i}$$

Where:

$\boldsymbol{Metabolite_i}$ is the metabolite concentration of i^{th} animal

$\boldsymbol{MY}_i$ is the month and the year that blood sample was taken from i^{th} animal

$\boldsymbol{AFC}_i$ is age at first calving of i^{th} animal

$\boldsymbol{Parity}_i$ is lactation number of i^{th} animal

$\boldsymbol{Health}_i$ is the health status of i^{th} animal

$\boldsymbol{Ration}_i$ is the type of fed ration of i^{th} animal

$\boldsymbol{RFI_i}$ is the group of i^{th} animal (High RFI or Low RFI).

$\boldsymbol{e_i}$ is the residual error Linear contrast among predictions for RFI groups was calculated using lsmeans package (Lenth, 2016). Data were reported as least square means with related standard errors. We declared significance at $P \leq 0.05$ and tendency was reported at $0.05 \leq$

$P \leq 0.10$. If any of independent variables except RFI were significant, the metabolite was adjusted to remove confounding effects. The adjusted metabolite concentrations were later used for multivariate analyses and RFI value predictions.

3.2.3 Multivariate analysis for classification

Adjusted metabolites were used for this analysis to make sure no confounding effects influence the result. First step was autoscaling of data which is the most popular method of data dispersion and was done according to (Xia, et al. 2013). The data points in this method were mean centered and divided by standard deviation.

After normalization process for all metabolites, statistical multivariate analysis for classification of animals was done using principal component analysis (PCA) partial least square discriminant analysis (PLS-DA). The permutation test on PLS-DA model was used for validation of model. PCA is an unsupervised clustering method that reduces space for high dimensional data set. It examines the variation in the data and reduces the complexity of data for interpretation. Scores plot were used to show similarities and differences between samples from high and low RFI groups. The first two axes of score and loading plot often is used to show the classification between groups of samples with PCA.

PLS-DA is a supervised learning algorithm that maximizes the covariance between response variable and features (metabolites). Characterization of data pattern is done with first two principal components of score plots. Important metabolites were sorted based on variable of importance projection (VIP). If the VIP for a metabolite is greater than 1.0, it suggests that the metabolite is statistically significant in classification of high and low RFI samples. All

multivariate analysis for classification of high and low RFI animals using metabolite profile was done using MetaboAnalyst (Xia et al. 2013)

Statistical significance of PLS-DA model was assessed with permutation testing. This test is done to make sure the discrimination of two classes on score plot and the value of diagnostic statistics are not gained by chance because of lucky choice of samples. Permutation test assumes that there is no difference between two groups that are randomly labeled by permutation. This test is an effort using existing data to ensure that the outcome of developed PLS-DA model is reproducible (Szymańska et al. 2012).

The performance of PLS-DA model in classification of high and low RFI animals is illustrated using receiver operating characteristic (ROC) curve. The ROC curve is made by plotting true positive rate (TPR) against false positive rate (FPR) at different threshold settings. The threshold is made by the classifier model when classification is done by more than one variable. TPR is the ratio of positive cases that are predicted positive correctly to all individuals that are predicted as positive. FPR is the ratio of cases that are predicted positive falsely to all individuals that predicted negative. In machine learning terminology, TPR and FPR are known as sensitivity and (1-Specificity). The sensitivity is the proportion of positives that are identified correctly. The specificity is the proportion of negatives that are identified correctly. Formula of these measures are as follow:

Sensitivity = TPR = $\frac{TP}{P}$, FPR = $\frac{FP}{N}$ = 1 - specificity (Yerushalmy, 1947)

3.2.4 Multiple linear regression for RFI value prediction

Multiple linear regression is the extended version of ordinary least square regression with multiple independent variables. This model considers the linear relationship of response variable

with predictors. Limitation of this model is that number of predictors cannot be larger than number of experimental units. The assumptions of normality of residuals and homogenous variance, and independence of residual were checked with appropriate tests and diagnostic plots in R (R-Core-Team, 2019).

Relative importance of variables in the model was estimated using relaimpo package which is one of the state-of-the-art methods (Grömping, 2006). The stability of estimations of the metabolites relative importance values were also computed using bootstrap technique of relaimpo package. Leave-one-out cross validation (LOOCV) of linear models were done using caret package (Kuhn 2005).

3.3 Results

A total of 33 serum metabolites were identified and quantified at 50, 150 and 240 DIM and comparisons of each metabolite between high and low RFI groups was performed using univariate linear regression analysis. Moreover, multivariate analysis was conducted to see if high and low RFI cows cluster separately and facilitate biomarker identification. Using multiple linear regression analyses panels of metabolites that can predict individual RFI phenotypes were identified at each time point during the lactation period. In the comparisons of each metabolite concentration between high and low RFI groups at 50, 150 and 240 Days in milk (DIM), fixed effects such as parity (P), age at first calving (AFC), sampling month and year (MY), health status (HS) and ration (R) were adjusted whenever significant ($P < 0.05$).

3.3.1 Univariate metabolite analysis

3.3.1.1 Serum metabolites affected by RFI category at early stage of lactation

At the early stage of lactation (50 DIM), 6 metabolites (glycerol, urea, creatinine and dimethyl sulfone) were significantly ($P<0.05$) different between high and low RFI groups. Low RFI cows had 37% more ($P = 0.00000166$) glycerol than high RFI cows (table 3.1). The narrow standard error of estimates of glycerol for both groups is considerable which is the reason for very small P-values at sample size of 20 per group of RFI. Serum concentration of glycerol was not affected ($P > 0.05$) by any of the fixed effects considered in the analysis. Urea has higher concentration in low RFI cows than high RFI group ($P = 0.000124$). The month and year of sampling had an effect on urea concentration. Elevated concentration of creatinine was observed in low RFI cows ($P = 0.014$) compared to the high RFI groups. All the fixed effects considered in the analysis had no effect ($P > 0.05$) on the concentration of this metabolite. Dimethyl sulfone had lower concentration in low RFI cows than high RFI ($P = 0.018$). Month and year of blood sampling time had also significant effect on dimethyl sulfone concentration.

3.3.1.2 Differences in serum metabolites concentrations between high and low RFI cows at mid lactation stage

The comparison of high and low RFI groups at the mid lactation (150 DIM) stage revealed that the serum concentrations of creatinine, glycerol, L-ornithine and L-lysine were significantly altered (table 3.2). The concentration of creatinine was increased ($P < 0.05$) in the low RFI compared to the high RFI cows. Parity, age at first calving and health had significant effect on the serum concentration of creatinine. Glycerol was the second most significant metabolite with higher concentration in low RFI cows. L-Ornithine had lower ($P < 0.05$) concentration in low

RFI cows compared to high RFI. L-ornithine concentration was significantly affected by month and year of sampling. L-Lysine had higher ($P < 0.05$) concentration in high RFI cows. It should be mentioned that parity effect was significant ($P < 0.05$) in concentration of L-lysine in blood serum in 150 DIM. Acetone had tendency of significance ($P=0.09$) with more concentration in low RFI animals. Our result showed that age at first calving and month-year of sampling affects the serum concentration of acetone in high and low RFI animal groups of lactating dairy Holstein cows.

3.3.1.3 Differences in metabolites concentration between high and low RFI groups at late lactation stage

A total of 8 metabolites (glycerol, acetone, citric acid, 3-hydroxybutyric acid, choline, creatinine, glycine, formate) significantly differed between high and low RFI cows**,** whereas isoleucine L-lysine and urea tended to differ between the RFI groups at the late lactation stage (240 DIM) (table 3.3). Interestingly, none of the fixed effects considered in the analyses had significant effect on the serum concentration of glycerol in lactating dairy cows at 240 DIM. Citric acid and acetone had higher ($P < 0.05$) concentration in low RFI cows than high RFI group. Month and year of blood sampling also had influence on concentration of acetone and citric acid. Creatinine, glycine, formate and 3-hydroxybutyric acid had higher concentration in low RFI animals in late lactation stage with no significant fixed effects. Choline showed lower concentration in serum of low RFI group than high RFI with significant effect of health status ($P<0.05$).

L-lysine and urea tended to increase in high RFI compared to the low RFI groups, whereas isoleucine tended to decrease in the high RFI group. Age at first calving and ration type had impact on concentration of isoleucine in 240 DIM. L-lysine concentration was only affected by

ration type, with a higher concentration in high RFI animals. Urea concentration was affected by month-year.

3.3.2 Multivariate analysis

3.3.2.1 Early lactation

Principal component analysis (PCA) did not show separation well between groups of RFI. Partial least squares discriminant analysis (PLS-DA) at the early lactation stage revealed that the first and second components of score plot distinguished low and high RFI dairy cows (Figure 3.1.a). A significant (P = 0,002) PLS-DA model was observed at early lactation stage indicating the separation observed between the two groups (low vs. high RFI) was very unlikely to be due to chance (Figure 3.1.c). The variable of importance in projection (VIP) score plot showed that glycerol, urea, creatinine, dimethyl sulfone, L-glutamine were the top 5 metabolites responsible for the separation of high and low RFI groups (Figure 3.1.b). The performance of top two metabolites (glycerol and urea) in classification of lactating cows to high and low RFI is shown in ROC curve (Figure 3.1.d). The area under the curve (AUC) of ROC plot was 0.936 with 95% confidence interval of (0.816 – 1.0).

3.3.2.2 Mid lactation

The partial least squares discriminant analysis at the mid lactation stage (150 DIM) showed a partly separation of the high and low RFI groups with tendency to be significant in permutation test (P = 0.09) (Figure 3.2 a and c). The VIP plot was used to rank the metabolites based on their importance in discriminating the high and low RFI cows. Creatinine, glycerol, L-lysine, L-histidine, L-ornithine, Formate, and L-Proline are the most powerful group discriminators with

VIP >1.0 (Figure 3.2 b). The AUC of ROC was 0.866 using top three metabolites (creatinine, glycerol and L-Lysine) in 150 DIM (Figure 3.2 d).

3.3.2.3 Late lactation stage

Multivariate classification analysis of high and low RFI dairy cows using day 240 of lactation blood samples with PLS-DA model was validated (figure 3.3.a). Permutation test had very small P-value of 0.002 and model was considered significant (figure 3.3.c). Important features of the model (VIP > 1.0) were respectively glycerol, citric acid, creatinine, acetone, beta hydroxy-butyric acid, formate, choline, L-lysine, isoleucine and urea (figure 3.3.b). The performance of top two metabolites (glycerol and creatinine) in classification of high and low RFI cows is shown with ROC curve (figure 3.3.d). The AUC of these two classifiers was 0.997 with confidence interval of (0.969 - 1.0).

3.3.3 Regression analysis for RFI value prediction

The average of RFI (Mcal/day) from 3 to 240 DIM was used as response variable and metabolites concentrations (µM) adjusted for significant fixed effects were used as predictor of RFI variables. The model at 50 DIM of lactation explained 71.76% of variation in RFI. Relative importance of metabolites is shown in Table 3.4. Regression coefficient of metabolites with standard error is provided in this table which shows with one-unit change of metabolite concentration (µM) of serum, how much RFI value would change and in which direction. Relative contribution to RFI could differ even if this experiment is repeated with identical conditions. Therefore, it is informative to have an estimate of the stability of relative importance of metabolites for prediction of RFI. Figure 3.5 shows the relative contribution of metabolites

with 95% confidence interval. The confidence interval is gained by bootstrap technique. This stability estimate means if we repeat the experiment 100 times with same conditions, in 95 times of that the result for relative contribution of metabolites will be in the range of the bar plot shown for each metabolite (Figure 3.5).

RFI was predicted from a set of 10 serum metabolites at the mid lactation stage (150 DIM) with stepwise multiple linear regression model (Table 3.5). The prediction model explained 73.74% of variation in RFI (Table 3.5). The relative contribution of metabolites from largest to smallest were creatinine, L-ornithine, glycerol, L-phenyl alanine, L-proline, L-threonine, citric acid, acetone, L-glutamine, and 3-hydroxy butyric acid. Among this group of metabolites, creatinine and L-ornithine had highest contribution to RFI prediction model. The stability of metabolite relative contribution to RFI is presented as a bar plot in figure 3.7.

At the late lactation stage (240 DIM), RFI could be predicted from 6 metabolites with model accuracy of 0.85. The linear regression model for 240 DIM could explain 85.21% of variation with 6 metabolites (table 3.6). Important metabolites in order were glycerol, L-alanine, pyruvic acid, creatinine, creatine, and formate. More than half of contribution of RFI in developed model for this sampling day can be explained with glycerol. Stability of 6 important metabolites of 240 DIM as predictors of RFI is shown in bar plot (figure 3.6)

Comparison of developed models for prediction of RFI is provided in a table form with each row representing a feature for model for comparison and columns as our results from early, mid and late lactation stages (Table 3.7). Multiple R^2 is the amount of variation in RFI that could be explained using the trained model in this experiment. This model feature had highest value for 240 DIM with 85.21 %. Adjusted R^2 is the penalized Multiple R^2 for number of variables in the model. Adjusted R^2 is the generally reported measure in literature. This model feature was best

in 240 DIM with 0.82 adjusted R^2. To avoid overfitting leave-one-out cross validation statistics of a model is reported. The purpose of all cross-validation measures is to show how adequate is the model for unseen data. Cross validated R^2 shows how much of unknown RFI values can be explained by metabolites. Cross validated R^2 was 0.71 for 240 DIM with highest value among other days. Root mean squared error (RMSE) and mean absolute error (MAE) are two most common measures of accuracy of model that represent the bias of predicted values from observed values. The lowest values for RMSE and MAE are respectively 1.50 and 1.23 from 240 DIM. Based on parsimony rule in statistics, simpler models with a smaller number of metabolites that can explain more variation of the response variable are better. Number of biomarkers for each model is provided in (table 3.7). The model for 240 DIM has also the least number of biomarkers in comparison to previous models made by 50 DIM and 150 DIM metabolite profile.

3.4 Discussion

Metabolomic profiling of dairy cows for classification of RFI groups is scarce (Wang & Kadarmideen, 2019). One of objectives of this study was to investigate if high and low RFI cows differ in serum metabolite concentrations. This was tested with a linear regression that included RFI class of animals and other factors that could potentially change metabolite concentration of serum. Three time points (50, 150 and 240 DIM) during the lactation period were selected to collect blood samples. These three days of sampling were selected in the middle of early, mid and late lactation stages as they are representative of metabolism status for lactating cows as it was explained in introduction (section 3.1.1). Univariate analysis could show us the effect of grouping of animals by high and low RFI on concentration of metabolites that were above limit of quantification (LOQ) on NMR platform. In univariate analysis, each metabolite is considered

separately. Therefore, the importance of metabolite in dairy cow health and production, the role of the metabolite in metabolism and the probable cause of significant different level of metabolite concentration are discussed here based on available evidence and knowledge of metabolites from the literature up to now. Multivariate and regression analyses which considered all adjusted metabolites of NMR platform together in a single model is also discussed in section 3.4.10. The agreement or disagreement of our result with similar studies in the literature is mentioned in each subsection of discussion.

3.4.1 Glycerol

The results revealed that the concentration of glycerol is consistently higher in low RFI than high RFI cows at all the three time points during the lactation period. Glycerol has a strong influence in dairy cow nutrition supplementation since decades ago (Kholif et al., 2019). Although the effect of glycerol on milk production has been different between studies, the interest for its use in dairy cows has increased. The reason is its effect when dairy cows are in negative energy balance. Increased blood glucose and decreased ketone bodies have been observed in many studies where glycerol was added to the diet (Osman et al. 2008; Paiva et al. 2016; Thoh, et al. 2017; Porcu et al. 2018; Saleem, et al. 2018). There is a high interest in dairy nutrition for the use of glycerol in diet at 5% to 8% level (Donkin, 2008). Despite our experiment did not have any nutritional treatment, our result suggests that low RFI animals naturally have higher glycerol concentration which is promising as it is helpful in providing energy for maintenance and production requirements.

The liver is responsible for synthesizing glycerol from adipose triglycerides. To obtain energy from fat by lipolysis, triglyceride breaks down to glycerol and three free fatty acids (FFA). Cells

from other tissues take up FFA after its release into the blood. The cell utilizes FFA in beta oxidation process and produces acetyl-CoA. Acetyl-CoA will subsequently lead to generation of adenosine three phosphate (ATP) in TCA cycle. The ATP provides energy for many biochemical processes in the cell (Wood, 2000).

The other end of triglyceride molecule is glycerol which enters the glycolytic pathway as dihydroxyacetone phosphate (DHPA). The final product of glycolysis is pyruvate. The pyruvate molecule will be processed by pyruvate oxidation, TCA cycle and finally oxidative phosphorylation to produce ATP (Wood, 2000). Glycerol is a precursor of glucose and its metabolic pathway is similar to glucose (Fisher et al. 1973). This metabolite was the most important metabolite in 50 and 240 DIM and was second most important metabolite in 150 DIM in univariate analysis. Our result suggests higher lipid metabolism in low RFI cows in comparison to high RFI cows.

3.4.2 Creatinine

Creatinine was the second most significant ($P<0.05$) metabolite that had different level of concentration between RFI groups in all sampling days. The concentration of creatinine showed persistent increase in low RFI in comparison to high RFI cows at all the three time points. Creatinine was not found to be influenced by other factors at 50 and 240 DIM, but this was not the case at 150 DIM. At this point (150 DIM), creatinine was significantly affected by parity, age at first calving and health status. Karisa et al. (2014) also reported association between RFI group and creatinine in beef cattle. Montanholi et al. (2017) explored the relationship between liver morphometry and blood metabolic profiles and reported higher concentration of creatinine in feed efficient animals which is in accordance with our results. Because they took liver samples

after slaughter of high and low RFI steers to generate photo microscopic images, histological quantification provided strong visual evidence of enlarged hepatocytes in low RFI animals. Larger hepatocytes area in low RFI steers from different regions of liver (portal triad ($P = 0.001$) and central vein ($P = 0.03$)) was accompanied by significant different ($P = 0.03$) level of plasma creatinine concentration. They concluded that low RFI animals have increased metabolism by liver. Creatinine was also found as one of the important variables of classification of high and low RFI of Nellore cattle (Consolo et al., 2020). The concentration of creatinine in liver and muscle tissues respectively were reported 0.44±0.16 and 1.71±0.34 μM for low RFI and 0.23±0.04 and 0.25±0.02 μM for high RFI beef cattle (Consolo et al., 2020). This agrees with our result that represents consistent significant difference ($P<0.05$) of serum creatinine concentration in lactating Holsteins in three days of blood sampling. The concentration of creatinine in low RFI animals is greater than high RFI group in serum samples of lactating Holsteins in our study which is also the case in liver and muscle samples of beef cattle. Creatinine is the converted form of creatine that is made in muscle. Creatinine, then leaves muscle through the blood stream and is carried to kidney. Finally, creatinine is excreted into the urine.

3.4.3 Urea

Urea was shown to be significantly different between low and high RFI groups at 50 DIM and there was a tendency ($0.05<P<0.10$) for it to be different at 240 DIM. Only month-year factor of blood sampling was also significant ($P<0.05$) for this metabolite at these days. This is in accordance with a study that evaluated the effect season on urea in dairy cows. The concentration of urea in blood was found to be significantly ($P<0.05$) lower in summer than spring and autumn

(Alberghina et al., 2013). Metabolism of urea, along with methane, was reported to be associated with RFI value in beef cattle (Karisa et al. 2014). In addition, urea was reported to have 4 times the concentration in muscle tissue of beef cattle in a low RFI group in comparison with high RFI (Consolo et al., 2020) which is in agreement with the results obtained in this study. Most excreted nitrogen in mammals is in the form of urea.

The urea is produced by nitrogen metabolism in the mitochondrial matrix of liver cells in a series of reaction that are known as urea cycle. The main objective of the urea cycle is to remove highly toxic ammonia from the body (Watford, 2003). The urea cycle starts with carbamoyl phosphate which is converted form of ammonia. In first step of urea cycle, carbamoyl phosphate needs ornithine to be converted to citrulline. This is where ornithine plays a direct role in the urea cycle (Wood, 2000).

The other amino acid is aspartate that has a role in the second step of the urea cycle, a reaction that consumes 3 ATP. Aspartate is made by transamination of glutamate. Lysine is an important precursor of glutamate and it has indirect impact on the urea cycle through glutamate and consequently aspartate (Wood, 2000).

Lysine and ornithine were reported as differentially expressed amino acids that are indirectly and directly related to urea cycle in a metabolomics study of growing heifers (Smeding et al. 2019). These authors suggest that down regulation of urea cycle occurs in low RFI heifers in comparison to high RFI (Smeding et al. 2019). This agrees with our result that shows ornithine and lysine have significantly ($P<0.05$) lower concentration in low RFI animals in 150 DIM. Furthermore, the result from 240 DIM represent a tendency of significance ($0.05<P<0.10$) for urea and lysine. The down regulation of amino acids that are related to urea cycle can be because of lower DMI of low RFI animals that is the cause of lower protein intake as it is suggested by

Smeding et al. (2019). This conclusion stays in the line with another study result that shows significantly (P=0.04) lower nitrogen in milk and feces of low RFI lactating Holsteins relative to high RFI (Rius et al., 2012

3.4.4 3-hydroxybutyric acid

The 3-hydroxybutyric acid (other name is beta hydroxybutyric acid) is a ketone body synthesized from fatty acids in liver. In addition to playing a role as an essential carrier of energy when glucose supply is low, beta hydroxybutyric acid was recently found to have an important role in cell signalling. The signalling function of this intermediate metabolite occurs both at the surface of cells and intercellularly to impact gene expression, lipid metabolism, metabolic rate and neural function (Newman & Verdin, 2017). Beta hydroxybutyric acid (BHB) is associated with higher genetic merit for milk fat yield in lactating dairy cows (Dechow, Baumrucker, Bruckmaier, & Blum, 2017). This metabolite was significantly ($P<0.001$) different between RFI groups in beef cattle but in the opposite direction. Positive correlation of RFI and beta hydroxybutyric acid is reported which means feed efficient beef cattle have less concentration of beta hydroxybutyric acid (Karisa et al. 2014). Our study found significantly ($P<0.05$) higher concentration of beta hydroxybutyric acid in low RFI animals at 240 DIM which contradicts the expression of this metabolite in beef cattle. This metabolite is known as one of the indicators of negative energy balance and elevated levels of BHB may have unfavorable effects on health and reproductive traits. But this is true for early lactation stage which is not the case in our result as we found the difference in late stage lactation when animal is increasing body weight. Consideration of blood metabolites while evaluating dairy cows for feed efficiency may help making decisions that maintain high level of health and reproductive traits (Dechow et al., 2017).

Future studies in dairy cow with focus on cell signalling role of beta hydroxybutyric acid may find the underlying reason of significant difference between high and low RFI animals.

3.4.5 Amino acids

Concentration of some amino acids was observed to be significantly different between RFI groups. Karisa et al. (2014) showed lysine as one of the metabolites that have tendency ($P = 0.09$) to be significantly different between high and low RFI groups of beef cattle. Wang and Kadarmideen (2019) reported lysine and isoleucine among other amino acids that are associated with RFI in dairy cows. The possible origin and biological reason for significantly different concentration of amino acids in the serum of RFI groups in this study is explained in relation to differences in other metabolites. For example, the role of lysine and ornithine in the section describing the urea cycle (3.4.3) and the importance of lysine, isoleucine and glycine in TCA cycle are explained in the section on citrate (3.4.8).

3.4.6 Acetone

Acetone was significantly ($P = 0.01$) different between RFI groups at 240 DIM. This metabolite has higher concentration in low RFI animals than high RFI group. Month and year of blood sampling affected significantly ($P<0.05$) the concentration of this metabolite. The pattern of higher concentration of acetone for low RFI animals is constant between day 150 and 240 of lactation results. There is not any evidence of different expression of acetone in high and low RFI lactating cows up to now, but acetone has been seen in a study of beef cattle that considered blood metabolomic biomarkers of divergent RFI groups (Karisa et al. 2014). This study had discovery and validation population to validate metabolite biomarkers. Acetone was significantly

associated with average body weight (ABW) of beef steers in both discovery and validation population (Karisa et al. 2014). In dairy cows, acetone concentration of milk and blood are indicators of ketosis but this happens in days around calving up to the end of early lactation stage around week 10 of lactation (van Knegsel et al., 2010). Our result did not show even a tendency of different acetone concentration in 50 DIM. There is a tendency of significant difference between high and low RFI groups in 150 DIM ($P = 0.09$) and this gap increases in 240 DIM ($P = 0.01$) which can show slight change of acetone concentration from mid-lactation to late lactation. Blood acetone in high levels is generally associated with ketosis, but this happens when dairy cow is in negative energy balance or the mammal is fasting or have caloric restrictions (Koeslag, et al. 1980). Glucose formation from acetone is also a minor way of metabolism in mammals (Kalapos, 2003). This form of conversion happens when there is small amount of acetone in blood which are made from fatty acids in blood. These fatty acids are mobilized and released from adipose tissue (Kalapos 2003). The difference in the regulation of the pathway that are involved in conversion of acetone to glucose might be the reason of higher acetone concentration in low RFI animal in mid and late lactation when there is no feed restriction or negative energy balance. The low concentration of acetone in mid (33.7 ± 6.39 μM) and late (57.20 ± 10.10 μM) lactation in our result in comparison to blood acetone level in cows that have ketosis (300-350 μM) is considerable (Belay, et al. 2017). This gap strengthens the possibility that acetone to glucose conversion is the biological reason of higher acetone level in low RFI group than high RFI in 150 and 240 DIM.

3.4.7 Formate

Formate is an intermediate metabolite in the rumen that is a source of carbon and hydrogen for rumen microbial population growth. Almost 18% of rumen methane (CH_4) is produced from formate. Microbes use the intercellular formate in rumen and convert it to methane. This process is called methanogenesis and one of the important contributors to methane production from formate in rumen is *Methanobacterium ruminantium* (Hungate et al. 1970). This microbe tends to have higher abundance (P=0.22) in high RFI lactating Holsteins in early stage of lactation than low RFI. Although this difference was not significant using 8 cows per group in that study, further investigation of abundance of other methanogenic bacteria plus this species together may make the role of rumen and its byproduct metabolites in feed efficiency clearer (Rius et al., 2012). Nevertheless, our result states formate has significantly (P=0.04) higher concentration in serum of low RFI cows. Recently, a study on dairy heifers reported greater CH_4 (g/Kg of DMI) in low RFI animals (Flay et al., 2019). This agrees partially with our result that shows formate which is the intermediate metabolite for methane (CH_4) production by microbes has higher concentration in serum. It is important to mention that CH_4/DMI is a measure that is adjusted for DMI, this measure of methane is independent from the argument that low RFI animals make less methane because of less DMI. Beef studies on methane have shown the same relationship between RFI class of animal and CH_4 emission. They showed low RFI beef cows produce 3.7% less CH_4 daily but have 9.9% higher CH_4 when related to DMI (g/Kg of DMI) (Manafiazar et al. 2020).

There are few reports of this metabolite in cattle, but it is believed this metabolite is a by-product from the rumen. The difference in concentration level of serum samples of high and low RFI cows might be because of differences in microbial activities in ruminal digestion of feed.

3.4.8 Citrate

Citrate (other name citric acid) has an important role in citric acid cycle (TCA cycle). Citrate is consumed by this cycle and regenerates with TCA cycle again. TCA cycle works closely with oxidative phosphorylation pathway to release energy in form of ATP from carbohydrates, fat, and protein. These biological processes take place in matrix of mitochondria. Citrate concentration was significantly higher (P = 0.01) in low RFI cows at 240 DIM. This is in accordance with the result of Karisa et al. (2014) who reported citrate as significantly different metabolite between high and low RFI animals. The higher concentration of citric acid as initiator of TCA cycle might be associated with upregulation of this pathway. This is supported by the finding that other metabolites associated with the TCA cycle are different in low RFI cows such as choline (P = 0.02) isoleucine (P = 0.07), lysine (P = 0.08) and glycine (P = 0.04). Lysine is one of the amino acids that is needed to make acetoacetyl - CoA then it converts to acetyl – CoA and finally citrate. Isoleucine is another amino acid necessary for acetyl CoA synthesis. Choline can be converted to acetyl choline and consequently acetyl-CoA and therefore play a role in the TCA cycle. Glycine has two ways to influence TCA cycle as it is needed to make pyruvate and pyruvate can be converted to oxaloacetate which is a main component of TCA cycle or pyruvate participates in synthesis of acetyl CoA and finally citrate (Wood, 2000).

In late lactation stage, cows are rebuilding body reserves that they have lost in early lactation stage. The difference between high and low RFI groups may be partly in regulation of pathways to rebuild muscle and fat tissues. Recent studies have found evidence that TCA cycle intermediates have other roles in addition to releasing energy such as DNA methylation, immunity and signalling molecules with functions in chromatin modifications. RFI is known as

one of complex traits that 20% of its variation is in additive genetic control (Manafiazar et al. 2015). Beef cattle studies in RFI has found differential pattern of DNA methylation among animals (Betancourt et al. 2016). Further studies using other biological data may make the biological picture of RFI regulation clearer in dairy cows as it has been done in beef cattle.

3.4.9 Dimethyl sulfone

The metabolite with largest P-value in univariate analysis of 50 DIM was dimethyl sulfone (Table 3.1). There is little information about this metabolite. In normal health status of humans, this metabolite exists in blood plasma and has a dietary source (Wishart et al. 2013). However gut microbial source has also mentioned in human metabolite literature (Maher et al., 2013). High correlation (r=0.69, P=10^{-9}) using statistical heterospectroscopy method is recently reported between dimethyl sulfone concentration in blood plasma and milk of Holsteins. Authors have assigned dimethyl sulfone to microbial products of rumen (Maher et al., 2013). Our study result shows this metabolite significantly (P = 0.01) higher in blood serum of high RFI cows. Future studies with consideration of microbial population structure and abundance of rumen, plus rumen and blood metabolites may be able to show evidence of dimethyl sulfone role and origin as in biological regulation of RFI in dairy cows.

3.4.10 Choline

Choline was significantly (P = 0.02) different between high and low RFI groups only in 240 DIM. This metabolite had less concentration in low RFI cows than high RFI. Consolo et al (2020) assigned the significantly different levels of choline in muscle samples of beef cattle to differences in energy metabolism of carbohydrate and fat-related compounds. Choline also

showed significant difference (P=0.008) between blood plasma samples of high and low RFI beef cattle (Karisa et al. 2014).

3.4.11 Multivariate analysis discussion

Multivariate analysis uses multiple metabolites at the same time as predictors of RFI class of dairy cows. Multivariate classification using PLS-DA captured 16.7% of variation of 33 metabolites in 50 DIM and showed fair discrimination. Glycerol, urea, creatinine, and dimethyl sulfone are respectively the most important variables in multivariate classification analysis. The order of VIP scores (Figure 3.1) from PLS-DA in 50 DIM matches exactly with order of sorted significant metabolites (Table 3.1) by P-values from univariate analysis. Two independent analysis showed same metabolites are important for classification of dairy cows for high and low RFI. In addition to glycerol, urea, creatinine, and dimethyl sulfone at 50 DIM, there are other metabolites with (VIP >1). They are L-glutamine, tyrosine and 3-hydroxy butyric acid that stands after first four metabolites with small difference from VIP score of one. Results from 150 DIM (Table 3.2) represented creatinine, glycerol, L-ornithine and L-lysine among important metabolites in the classification model. The univariate analysis results (Table 3.2) also showed the same metabolites as significant or tendency of significance at this time point during lactation. Glycerol was the most important metabolite in VIP score of 240 DIM with score of 3.5 which was considerable in comparison to citric acid, creatinine, acetone, 3-hydroxy butyric acid, formate, choline, lysine, isoleucine, and urea with VIP score in range 1 to 1.7.

Score plots provided by PLS-DA model are projected metabolite concentrations on a hyperplane that maximized the covariance of metabolites concentrations and class of RFI. The higher scores show more importance of the variable in separation of RFI classes which is called variable of

importance projection (VIP) score (Worley & Powers, 2013). Visual separation of RFI groups by score plots of PLS-DA showed the best discrimination in 240 DIM (Figure 3.3) was gained and after that 50 DIM (Figure 3.1) where the ellipses had less common area than 150 DIM (Figure 3.2). All animals were sorted based on the average of cumulative RFI in all 240 days. Thus, prediction of whether the lactating cow belongs to high or low RFI group at day 50 which is 190 days before end of experiment with metabolite profile is reasonable to have less accuracy than using day 240 metabolite profile. As the metabolism status of cows at day 240 is partly a result of the past metabolism rhythm at day 50 and 150, however this relationship is not only dependent on time because day 50 (Figure 3.1) showed better separation of RFI classes than day 150 (Figure 3.2). Day 50 is in the vicinity of peak of production and peak of lactation for most cows and has unique importance in dairy cow production and management (Amaral-Phillips, 2012). Milk yield of 305-day lactation is tightly related to peak of milk production so that there are models that can predict overall milk production of animal using peak of milk yield with decent accuracy (R-squared = 0.69) (Mellado et al. 2011, Amaral-Phillips 2012). The production and health of animal is a function of metabolism and feed efficiency have same relationship (Rius et al., 2012). The high accuracy of our result for prediction of RFI in 240 days using day 50 serum metabolite profile may be another indication of the importance of metabolism status in peak of production on the performance of animal in overall lactation period.

Classification of high and low RFI beef cattle using muscle metabolites with PLS-DA model showed glycine, choline, creatinine, 3-hydroxybutyric acid, and glycerol with VIP over 1.0 which are common with at least one of our sampling days VIP score results. Glycerol and creatinine had high VIP consistently in our three blood sampling days (Consolo et al., 2020).

This study did same analysis using liver samples too and VIP score of metabolites showed citric acid, creatinine, histidine and urea common with at least one of our sampling days.

This experiment to best of our knowledge was the first one that focused on biological differences in feed efficiency of Holsteins by stage of lactation. For example Wang et al (2019) did not mention the time of blood sampling in the report.

The other aspect is the larger sample size of animals in each group of RFI (n=20) in our study in comparison to other studies that tried to explain part of metabolic difference between high and low RFI animals. For example, metabolomics approach of blood plasma, immunoassay approach of blood serum concentrations and response of blood plasma hormones and metabolites to metabolic challenges had respectively 5, 14 and 9 animals per group of RFI (Wang and Kadarmideen 2019, Xi et al. 2016, Nürnberg, et al. 2019). This has led to lower statistical power which would result in smaller number of significant biomarkers and less certainty in making conclusions.

Moreover, our research in comparison to other reported papers is the length of time that animals are considered for RFI estimation which shows the efficiency of animal in feed and body resources utilization. Li et al. (2017) wrote a paper only about this issue as it was to some extent neglected. Researchers consider animals for RFI in time intervals that are different. However, Li et al. (2017) discussed around the differences of heritability of RFI along lactation and between primiparous and multiparous cows. This concept is applicable to underlying metabolic changes of RFI because genetic basis is part of the variation of RFI as they showed its significant heritability along lactation period. As an example the experiment that was done by Xi et al. (2016) used 50 days and RFI values are considered for that limited time period. Other studies such as Wang and Kadarmideen (2019) did not clearly mention the details of sampling day

whether it was done on early lactation or other stages. This information is essentially needed to make conclusion on metabolic efficiency of energy in dairy cows by stage of lactation and across lactations in future studies.

Metabolomics is useful to reveal biological differences between groups of individuals if confounding effects such as diet, age and other important variables (depends on experiment) would be considered (Xia and Wishart 2011, Emwas et al. 2015). The only confounding factor that Wang and Kadarmideen (2019) used in their univariate analysis of metabolites was parity. In our study, we considered month-year of blood sampling, age at first calving (AFC), ration type, health status and parity to remove confounding effects. The above-mentioned variables that we considered were identical to those that were used as fixed effects in calculation of RFI component traits using random regression model. Because it is believed that metabolites build the metabolic foundation of RFI based on researchers' experience in beef cattle studies (Karisa et al. 2014, Consolo et al., 2020). Exclusion of other variables that have significant effect on metabolite concentration can either reduce or increase the classification result such as P-value and least square means estimates of metabolite for high and low RFI groups. If there is no significant effect of these variables, the metabolite concentration would not be adjusted. This step is necessary to avoid misleading result and consequently unreproducible conclusions.

3.5 Conclusion

In conclusion, the results of this study suggest that metabolomics is useful for prediction of RFI in lactating dairy cows. Important metabolites that have potential predictive power were introduced for classification of animals in different stages of lactation. This study also provided regression prediction models for RFI values for the first time in dairy cows with accuracy

measures. As expected, other important variables that influence concentration of metabolites were detected. This knowledge is useful for design and implementation of experiments with larger sample size. Metabolites showed strong prediction power in this study. Given this strong relationship be reported in future studies, it would be worth to estimate heritability of metabolite biomarkers to be considered for genomic selection. We see considerable potential for reported biomarkers to contribute to this development.

3.6 References

Alberghina, D., Piccione, G., Casella, S., Panzera, M., Morgante, M., & Gianesella, M. (2013). The effect of the season on some blood metabolites and haptoglobin in dairy cows during postpartum period. *Archives Animal Breeding*, *56*(1), 354–359. https://doi.org/10.7482/0003-9438-56-035

Amaral-Phillips, D. M. (2012). *Using Peak and Summit Milk to Evaluate Your Dairy's Management Programs*. Retrieved from https://afs.ca.uky.edu/files/using_peak_and_summit_milk_to_evaluate_your_dairys_management_programs.pdf

Armbruster, D. A., & Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist. Reviews*, *29 Suppl 1*(Suppl 1), S49-52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18852857

Arthur, P. F., Archer, J. A., Johnston, D. J., Herd, R. M., Richardson, E. C., & Parnell, P. F. (2001). Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *Journal of Animal Science*, *79*(11), 2805–2811. https://doi.org/10.2527/2001.79112805x

Belay, T. K., Dagnachew, B. S., Kowalski, Z. M., & Ådnøy, T. (2017). An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. *Journal of Dairy Science*, *100*(8), 6312–6326. https://doi.org/10.3168/jds.2016-12252

Betancourt, F., Lachance, H., Friedman, S., Perlee, S., & McKay, S. (2016). Exploring DNA Methylation in the Bovine 5 ’ AMPK Gene Family (p. 1). Burlington: University of Vermont. Retrieved from https://www.researchgate.net/profile/Hannah_Lachance/publication/309121378_Exploring_DNA_Methylation_in_the_Bovine_5’_AMPK_Gene_Family/links/57ffc2e308ae6fc7fc6522d0/Exploring-DNA-Methylation-in-the-Bovine-5-AMPK-Gene-Family.pdf

Cantalapiedra-Hijar, G., Guarnido, P., Schiphorst, A. M., Robins, R. J., Renand, G., & Ortigues-Marty, I. (2019). Natural 15N abundance in specific amino acids indicates associations between transamination rates and residual feed intake in beef cattle. *Journal of Animal Science*, *98*(6), 1–7. https://doi.org/10.1093/jas/skaa171

Consolo, N. R. B., Buarque, V. L. M., Silva, J., Poleti, M. D., Barbosa, L. C. G. S., Higuera-Padilla, A., … Silva, S. L. (2020). Muscle and liver metabolomic signatures associated with residual feed intake in Nellore cattle. *Animal Feed Science and Technology*, *8401*(20). https://doi.org/10.1016/j.anifeedsci.2020.114757

Dechow, C. D., Baumrucker, C. R., Bruckmaier, R. M., & Blum, J. W. (2017). Blood plasma traits associated with genetic merit for feed utilization in Holstein cows. *Journal of Dairy Science*, *100*(10), 8232–8238. https://doi.org/10.3168/jds.2016-12502

Derno, M., Nürnberg, G., & Kuhla, B. (2019). Characterizing the metabotype and its persistency in lactating Holstein cows: An approach toward metabolic efficiency measures. *Journal of*

Dairy Science, *102*(7), 6559–6570. https://doi.org/10.3168/jds.2019-16274

Donkin, S. S. (2008). Glycerol from biodiesel production: The new corn for dairy cattle. *Revista Brasileira de Zootecnia*, *37*(SPECIALISSUE), 280–286. https://doi.org/10.1590/S1516-35982008001300032

Emwas, A. H., Luchinat, C., Turano, P., Tenori, L., Roy, R., Salek, R. M., … Wishart, D. S. (2015, November 21). Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. *Metabolomics*. Springer New York LLC. https://doi.org/10.1007/s11306-014-0746-7

Fisher, L. J., Erfle, J. D., Lodge, G. A., & Sauer, F. D. (1973). Effects of propylene glycol or glycerol supplementation on feed intake, milk yield and composition, and incidence of ketosis. *Canadian Journal of Animal Science*, *53*(2), 289–296. https://doi.org/10.4141/cjas73-045

Flay, H. E., Kuhn-Sherlock, B., Macdonald, K. A., Camara, M., Lopez-Villalobos, N., Donaghy, D. J., & Roche, J. R. (2019). Hot topic: Selecting cattle for low residual feed intake did not affect daily methane production but increased methane yield. *Journal of Dairy Science*, *102*(3), 2708–2713. https://doi.org/10.3168/jds.2018-15234

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., … Fukumasu, H. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-41813-x

French, P. (2017). *Applied Animal Nutrition Feedstuffs* (3rd Editio). Upper Saddle River, NJ, United States: Pearson Education (US). Retrieved from https://courses.ecampus.oregonstate.edu/ans312/seven/cows_6_trans.htm

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS ONE*, *12*(5), 1–26. https://doi.org/10.1371/journal.pone.0177675

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, *17*(1), 1–27. https://doi.org/10.18637/jss.v017.i01

Hailemariam, D., Mandal, R., Saleem, F., Dunn, S. M., Wishart, D. S., & Ametaj, B. N. (2014). Identification of predictive biomarkers of disease state in transition dairy cows. *Journal of Dairy Science*, *97*(5), 2680–2693. https://doi.org/10.3168/jds.2013-6803

Herd, R. M., & Arthur, P. F. (2009). Physiological basis for residual feed intake. *Journal of Animal Science*, *87*(14 Suppl), 64–71. https://doi.org/10.2527/jas.2008-1345

Hungate, R. E., Smith, W., Bauchop, T., Yu, I., & Rabinowitz, J. C. (1970). Formate as an intermediate in the bovine rumen fermentation. *Journal of Bacteriology*, *102*(2), 389–397. https://doi.org/10.1128/jb.102.2.389-397.1970

Jorge-Smeding, E., Renand, G., Centeno, D., Pétéra, M., Durand, S., Polakof, S., & Cantalapiedra-Hijar, G. (2019). Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. In *International Symposium on Energy and Protein Metabolism and Nutrition (ISEP)* (pp. 231–232). https://doi.org/10.3920/978-90-8686-891-9_50

Kadarmideen, H. N. (2008). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Kalapos, M. P. (2003, May 2). On the mammalian acetone metabolism: From chemistry to clinical implications. *Biochimica et Biophysica Acta - General Subjects*. Elsevier.

https://doi.org/10.1016/S0304-4165(03)00051-5

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., … Plastow, G. S. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livestock Science*, *165*(1), 200–211. https://doi.org/10.1016/j.livsci.2014.03.002

Karisa, B., Moore, S., & Plastow, G. (2013). Complicated Relationships: A Review of Biological Interaction Networks and Pathways in Animal Science. *Springer Science Reviews*, *1*(1–2), 73–83. https://doi.org/10.1007/s40362-013-0005-8

Karisa, B., Moore, S., & Plastow, G. (2014). Analysis of biological networks and biological pathways associated with residual feed intake in beef cattle. *Animal Science Journal*, *85*(4), 374–387. https://doi.org/10.1111/asj.12159

Karisa, Brian K. (2013). Candidate Genes , Metabolites and Biological Pathways Associated With Residual Feed Intake and Carcass. (Publication No. NS27583.) [PhD thesis, University of Alabama]. ProQuest Dissertations & ThesesGlobal

Kennedy, P.M & Milligan, L. P. (1980). the Degradation and Utilization of Endogenous Urea in the Gastrointestinal Tract of Ruminants: a Review. *Canadian Journal of Animal Science*, *60*(2), 205–221. https://doi.org/10.4141/cjas80-030

Khansefid, M., Millen, C. A., Chen, Y., Pryce, J. E., Chamberlain, A. J., Vander Jagt, C. J., … Goddard, M. E. (2017). Gene expression analysis of blood, liver, and muscle in cattle divergentlya selected for high and low residual feed intake. *Journal of Animal Science*, *95*(11), 4764–4775. https://doi.org/10.2527/jas2016.1320

Kholif, A. E. (2019). Glycerol use in dairy diets: A systemic review. *Animal Nutrition*, *5*(3), 209–216. https://doi.org/10.1016/j.aninu.2019.06.002

Koeslag, J. H., Noakes, T. D., & Sloan, A. W. (1980). Post-exercise ketosis. *The Journal of Physiology*, *301*(1), 79–90. https://doi.org/10.1113/jphysiol.1980.sp013190

Lenth, R. V. (2016). Least-Squares Means: The *R* Package lsmeans. *Journal of Statistical Software*, *69*(1). https://doi.org/10.18637/jss.v069.i01

Li, B., Berglund, B., Fikse, W. F., Lassen, J., Lidauer, M. H., Mäntysaari, P., & Løvendahl, P. (2017). Neglect of lactation stage leads to naive assessment of residual feed intake in dairy cattle. *Journal of Dairy Science*, *100*(11), 9076–9084. https://doi.org/10.3168/jds.2017-12775

Maher, A. D., Hayes, B., Cocks, B., Marett, L., Wales, W. J., & Rochfort, S. J. (2013). Latent biochemical relationships in the blood-milk metabolic axis of dairy cows revealed by statistical integration of 1H NMR spectroscopic data. *Journal of Proteome Research*, *12*(3), 1428–1435. https://doi.org/10.1021/pr301056q

Manafiazar, G., Baron, V. S., McKeown, L., Block, H., Ominski, K., Plastow, G., & Basarab, J. A. (2020). Methane and carbon dioxide emissions from yearling beef heifers and mature cows classified for residual feed intake under drylot conditions. *Canadian Journal of Animal Science*, *100*(3), 522–535. https://doi.org/10.1139/cjas-2019-0032

Manafiazar, G., Goonewardene, L., Miglior, F., Crews, D. H., Basarab, J. A., Okine, E., & Wang, Z. (2015). Genetic and phenotypic correlations among feed efficiency, production and selected conformation traits in dairy cows. *Animal*, *10*(3). https://doi.org/10.1017/S1751731115002281

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., & Wang, Z. (2012). Factors That Affect Energy Efficiency and Indirect Response of Selection for Efficiency on Related Traits. https://doi.org/http://dx.doi.org/10.5772/50776

Martin, O., & Sauvant, D. (2007). Dynamic model of the lactating dairy cow metabolism. *Animal*, *1*(8), 1143–1166. https://doi.org/10.1017/S1751731107000377

Kuhn, M. (2005). Building Predictive Models in R Using the caret Package. *Journal Of Statistical Software*, *14*(September), 981029. https://doi.org/10.18637/jss.v081.b02

Mellado, M., Antonio-Chirino, E., Meza-Herrera, C., Veliz, F. G., Arevalo, J. R., Mellado, J., & de Santiago, A. (2011). Effect of lactation number, year, and season of initiation of lactation on milk yield of cows hormonally induced into lactation and treated with recombinant bovine somatotropin. *Journal of Dairy Science*, *94*(9), 4524–4530. https://doi.org/10.3168/jds.2011-4152

Moe, P. W. (1981). Energy Metabolism of Dairy Cattle. *Journal of Dairy Science*, *64*(6), 1120–1139. https://doi.org/10.3168/jds.S0022-0302(81)82692-6

Montanholi, Y. R., Haas, L. S., Swanson, K. C., Coomber, B. L., Yamashiro, S., & Miller, S. P. (2017). Liver morphometrics and metabolic blood profile across divergent phenotypes for feed efficiency in the bovine. *Acta Veterinaria Scandinavica*, *59*(1), 1–11. https://doi.org/10.1186/s13028-017-0292-1

Mulligan, F., O'Grady, L., Rice, D., & Doherty, M. (2006). Production diseases of the transition cow: Body condition score and energy balance. *Irish Veterinary Journal*, *59*(9), 505–510.

Neves, R. C., Leno, B. M., Curler, M. D., Thomas, M. J., Overton, T. R., & McArt, J. A. A. (2018). Association of immediate postpartum plasma calcium concentration with early-lactation clinical diseases, culling, reproduction, and milk production in Holstein cows. *Journal of Dairy Science*, *101*(1), 547–555. https://doi.org/10.3168/jds.2017-13313

Newman, J. C., & Verdin, E. (2017). β-Hydroxybutyrate: A Signaling Metabolite. *Annual Review of Nutrition*, *37*, 51–76. https://doi.org/10.1146/annurev-nutr-071816-064916

Nkrumah, J. D. Okine, E. K., Mathison, G. W., Schmid, K., Li, C., Basarab, J. A., … Moore, S. S. (2006). Relationships of feedlot feed efficiency, performance, and feeding behaviour with metabolic rate, methane produciton, and energy partitioning in beef cattle. *Journal of Animal Science*, 84(1), 145–153. Retrieved from https://doi.org/10.2527/2006.841145x

O'Sullivan, M., Butler, S. T., Pierce, K. M., Crowe, M. A., O'Sullivan, K., Fitzgerald, R., & Buckley, F. (2020). Reproductive efficiency and survival of Holstein-Friesian cows of divergent Economic Breeding Index, evaluated under seasonal calving pasture-based management. *Journal of Dairy Science*, *103*(2), 1685–1700. https://doi.org/10.3168/jds.2019-17374

Osman, M. A., Allen, P. S., Mehyar, N. A., Bobe, G., Coetzee, J. F., Koehler, K. J., & Beitz, D. C. (2008). Acute metabolic responses of postpartal dairy cows to subcutaneous glucagon injections, oral glycerol, or both. *Journal of Dairy Science*, *91*(9), 3311–3322. https://doi.org/10.3168/jds.2008-0997

Porcu, C., Manca, C., Cabiddu, A., Dattena, M., Gallus, M., Pasciu, V., … Molle, G. (2018). Effects of short-term administration of a glucogenic mixture at mating on feed intake, metabolism, milk yield and reproductive performance of lactating dairy ewes. *Animal Feed Science and Technology*, *243*, 10–21. https://doi.org/10.1016/j.anifeedsci.2018.06.012

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., … Wishart, D. S. (2011). The human serum metabolome. *PLoS ONE*, *6*(2). https://doi.org/10.1371/journal.pone.0016957

R-Core-Team. (2019). R: A language and environment for statistical computing. Retrieved April 22, 2020, from https://www.r-project.org/

Ravanbakhsh, S., Liu, P., Bjordahl, T. C., Mandal, R., Grant, J. R., Wilson, M., … Wishart, D. S.

(2015). Accurate, fully-automated NMR spectral profiling for metabolomics. *PLoS ONE*, *10*(5), 1–15. https://doi.org/10.1371/journal.pone.0124219

Richardson, E. C., & Herd, R. M. B. (2004). Biological basis for variation in residual feed intake in beef cattle. 2. Synthesis of results following divergent selection Cooperative Research Centre for Cattle and Beef Quality . *Australian Journal of Experimental Agriculture*, *44*, 431–440. https://doi.org/10.1071/EA02221

Rius, A. G., Kittelmann, S., Macdonald, K. A., Waghorn, G. C., Janssen, P. H., & Sikkema, E. (2012). Nitrogen metabolism and rumen microbial enumeration in lactating cows with divergent residual feed intake fed high-digestibility pasture. *Journal of Dairy Science*, *95*(9), 5024–5034. https://doi.org/10.3168/jds.2012-5392

Saleem, A. M., Zanouny, A. I., & Singar, A. M. (2018). Effect of glycerol supplementation during early lactation on milk yield, milk composition, nutrient digestibility and blood metabolites of dairy buffaloes. *Animal*, *12*(4), 757–763. https://doi.org/10.1017/S175173111700180X

Salleh, M. S., Mazzoni, G., Höglund, J. K., Olijhoek, D. W., Lund, P., Løvendahl, P., & Kadarmideen, H. N. (2017). RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high- and low-residual feed intake in Nordic dairy cattle. *BMC Genomics*, *18*(1), 1–17. https://doi.org/10.1186/s12864-017-3622-9

Saude, E. J., Slupsky, C. M., & Sykes, B. D. (2006). Optimization of NMR analysis of biological fluids for quantitative accuracy. *Metabolomics*, *2*(3), 113–123. https://doi.org/10.1007/s11306-006-0023-5

Serkova, N. J., Standiford, T. J., & Stringer, K. A. (2011). The emerging field of quantitative

blood metabolomics for biomarker discovery in critical illnesses. *American Journal of Respiratory and Critical Care Medicine*, *184*(6), 647–655. https://doi.org/10.1164/rccm.201103-0474CI

Szymańska, E., Saccenti, E., Smilde, A. K., & Westerhuis, J. A. (2012). Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*, *8*(S1), 3–16. https://doi.org/10.1007/s11306-011-0330-3

Thévenot, E. A., Roux, A., Xu, Y., Ezan, E., & Junot, C. (2015). Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *Journal of Proteome Research*, *14*(8), 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

Thoh, D., Pakdeechanuan, P., & Chanjula, P. (2017). Effect of supplementary glycerin on milk composition and heat stability in dairy goats. *Asian-Australasian Journal of Animal Sciences*, *30*(12), 1711–1717. https://doi.org/10.5713/ajas.17.0066

van Knegsel, A. T. M., van der Drift, S. G. A., Horneman, M., de Roos, A. P. W., Kemp, B., & Graat, E. A. M. (2010). Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. *Journal of Dairy Science*, *93*(7), 3065–3069. https://doi.org/10.3168/jds.2009-2847

Wang, X., & Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites*, *9*(7), 1–16. https://doi.org/10.3390/metabo9070151

Wang, X., Han, W., Yang, J., Westaway, D., & Li, L. (2019). Development of chemical isotope labeling LC-MS for tissue metabolomics and its application for brain and liver metabolome

profiling in Alzheimer's disease mouse model. *Analytica Chimica Acta*, *1050*, 95–104. https://doi.org/10.1016/j.aca.2018.10.060

Watford, M. (2003). The urea cycle: Teaching intermediary metabolism in a physiological setting. *Biochemistry and Molecular Biology Education*, *31*(5), 289–297. https://doi.org/10.1002/bmb.2003.494031050249

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., … Scalbert, A. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research*, *41*(D1), 801–807. https://doi.org/10.1093/nar/gks1065

Wood, E. (2000). Lehninger's principles of biochemistry. *Biochemistry and Molecular Biology Education*, *28*(5), 286–287. https://doi.org/10.1016/s1470-8175(00)00020-5

Worley, B., & Powers, R. (2013). Multivariate Analysis in Metabolomics. *Current Metabolomics*, *1*(1), 92–107. https://doi.org/10.2174/2213235x11301010092

Xi, Y. M., Wu, F., Zhao, D. Q., Yang, Z., Li, L., Han, Z. Y., & Wang, G. L. (2016). Biological mechanisms related to differences in residual feed intake in dairy cows. *Animal*, *10*(8), 1311–1318. https://doi.org/10.1017/S1751731116000343

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, *9*(2), 280–299. https://doi.org/10.1007/s11306-012-0482-9

Xia, J., Sinelnikov, I. V., Han, B., & Wishart, D. S. (2015). MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Research*, *43*(W1), W251–W257. https://doi.org/10.1093/nar/gkv380

Yao, C., Spurlock, D. M., Armentano, L. E., Page, C. D., VandeHaar, M. J., Bickhart, D. M., & Weigel, K. A. (2013). Random Forests approach for identifying additive and epistatic single

nucleotide polymorphisms associated with residual feed intake in dairy cattle. *Journal of Dairy Science*, *96*(10), 6716–6729. https://doi.org/10.3168/jds.2012-6237

Yerushalmy, J. (1947). Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Reports*, *62*(40), 1432–1449. https://doi.org/10.2307/4586294

Yu, Z., Kastenmüller, G., He, Y., Belcredi, P., Möller, G., Prehn, C., … Wang-Sattler, R. (2011). Differences between human plasma and serum metabolite profiles. *PLoS ONE*, *6*(7), 1–6. https://doi.org/10.1371/journal.pone.0021230

Table 3.1 Least square means of metabolites that are significantly different between high and low RFI dairy cows at day 50 of lactation. Significant fixed effects are M: month-year, A: age at first calving, L: lactation, H: health status, R: ration.

Metabolite	High RFI		Low RFI		P-value	Fixed Effects	HMDBID
	Conc(μM)	SE	Conc(μM)	SE			
Glycerol	389.00	19.20	543.00	21.60	1.66E-06		HMDB0000131
Urea	1290.00	83.20	1784.00	81.20	1.24E-04	M	HMDB0000294
Creatinine	60.20	2.26	68.40	2.26	1.41E-02		HMDB0000562
Dimethyl sulfone	46.60	5.34	31.70	3.55	1.80E-02	M	HMDB0004983

Table 3.2 Least square means of metabolites that are significantly different between high and low RFI dairy cows in day 150 of lactation. Significant fixed effects are M: month-year, A: age at first calving, L: lactation, H: health status, R: ration.

Metabolite	High RFI		Low RFI		P-value	Fixed Effects	HMDBID
	LSM(μM)	SE	LSM(μM)	SE			
Creatinine	59.90	3.19	72.90	3.72	2.23E-03	L, A, H	HMDB0000562
Glycerol	286.00	12.60	331.00	12.60	1.43E-02		HMDB0000131
L-Ornithine	78.40	5.97	60.80	4.90	2.67E-02	M	HMDB0000214
L-Lysine	93.60	4.54	80.30	5.33	3.48E-02	L	HMDB0000182
Acetone	19.60	4.65	33.70	6.39	9.47E-02	M, A	HMDB0001659

Table 3.3 Least square means of metabolites that are significantly different between high and low RFI dairy cows in day 240 of lactation. Significant fixed effects are M: month-year, A: age at first calving, L: lactation, H: health status, R: ration.

Metabolite	High RFI		Low RFI		P-value	Fixed Effects	HMDBID
	LSM(µM)	SE	LSM(µM)	SE			
Glycerol	292.00	13.00	470.00	13.00	7.88E-12		HMDB0000131
Acetone	28.10	5.40	57.20	10.10	1.07E-02	M	HMDB0001659
Citric acid	190.00	10.55	227.00	9.64	1.14E-02	M	HMDB0000094
3-hydroxybutyric acid	557.00	77.10	938.00	129.90	1.32E-02		HMDB0000357
Choline	8.17	1.10	6.28	1.17	2.17E-02	H	HMDB0000097
Creatinine	65.00	2.29	72.40	2.29	3.32E-02		HMDB0000562
Glycine	273.00	17.70	329.00	24.70	4.06E-02		HMDB0000123
Formate	31.10	1.54	36.10	1.79	4.19E-02		HMDB0000142
Isoleucine	97.70	6.84	114.70	6.26	7.50E-02	A, R	HMDB0000172
L-Lysine	71.20	9.65	57.50	9.65	8.37E-02	R	HMDB0000182
Urea	1551.00	84.50	1344.00	103.00	9.95E-02	M	HMDB0000294

Table 3.4 Relative importance of metabolites in regression model to predict RFI values with metabolite profile at 50 DIM of lactation

Metabolite	Relative Contributions to R^2	Coefficient (SE)
Glycerol	0.22	-1.40 (0.34)
Urea	0.19	-1.19 (0.34)
Choline	0.08	-0.98 (0.32)
Dimethyl sulfone	0.06	0.98 (0.34)
Acetone	0.04	-1.23 (0.36)
Creatine	0.03	1.23 (0.34)
Isoleucine	0.03	0.80 (0.39)
L-Alanine	0.02	-0.71 (0.39)
Methionine	0.02	-0.85 (0.42)
D-Glucose	0.02	0.54 (0.31)
Model R^2	**0.7176**	**-**

.

Table 3.5 Relative importance of metabolites in regression model to predict RFI values with metabolite profile of day 150 of lactation.

Metabolite	Relative Contributions to R^2	Coefficient (SE)
Creatinine	0.31	-2.12 (0.36)
L-Ornithine	0.15	1.07 (0.35)
Glycerol	0.09	-1.09 (0.31)
L-Phenylalanine	0.05	0.85 (0.34)
L-Proline	0.04	1.26 (0.43)
L-Threonine	0.03	-1.16 (0.40)
Citric acid	0.03	1.01 (0.38)
Acetone	0.02	1.16 (0.48)
L-Glutamine	0.02	0.60 (0.31)
3-hydroxybutyric acid	0.01	-1.20 (0.46)
Model R^2	**0.7374**	**-**

Table 3.6 relative importance of metabolites in regression model to predict RFI values with metabolite profile of day 240 of lactation.

Metabolite	Relative Contributions to R^2	Coefficient (SE)
Glycerol	0.44	-1.96 (0.22)
L-Alanine	0.06	-1.03 (0.21)
Pyruvic acid	0.07	1.15 (0.22)
Creatinine	0.16	-1.15 (0.23)
Creatine	0.05	1.18 (0.23)
Formate	0.07	-0.52 (0.21)
Model R^2	**0.8521**	**-**

Table 3.7 Accuracy measures of regression model using metabolite profile form day 50, 150 and 240 of lactation to predict RFI for dairy lactating cows.

Model Feature	Metabolites as predictors		
	Day 50	**Day 150**	**Day 240**
Multiple R^2	0.7176	0.7374	0.8521
Adjusted R^2	0.6202	0.6469	0.8253
Cross validated R^2	0.5214	0.2477	0.7122
LOOCV RMSE	1.9377	2.7750	1.5096
LOOCV MAE	1.4740	2.2654	1.2309
Biomarkers	10	10	6

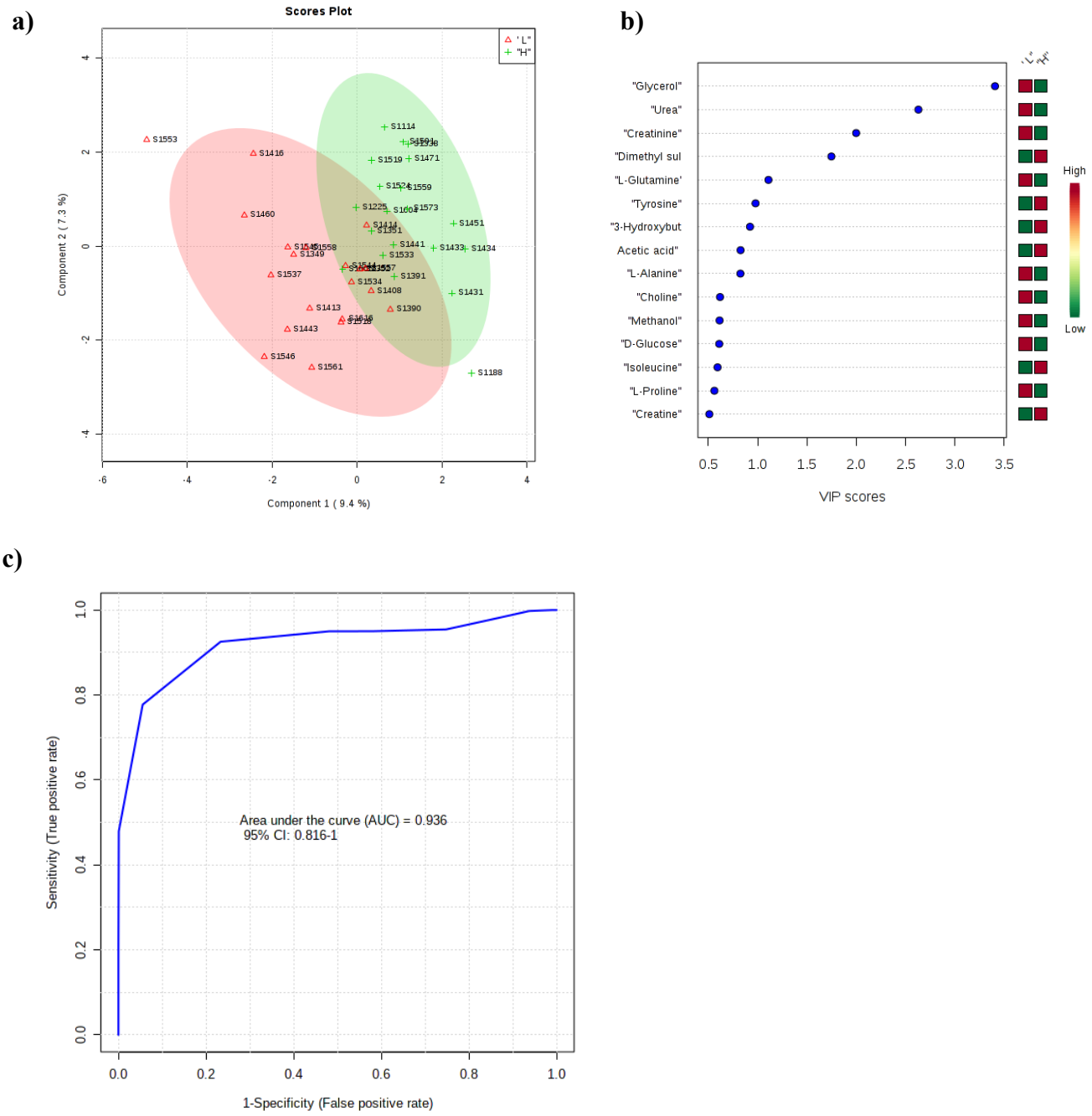

Figure 3.1 a) Score plots between first and second components of partial least square discriminant analysis (PLS-DA) on left panel and **b)** variable of importance in projection scores of PLS-DA analysis using metabolite profile of day 50 of lactation for classification of high and low RFI groups. **c)** The performance of top two serum metabolites (glycerol and urea) as biomarkers of RFI is presented in ROC curve. AUC = 0.936 (CI: 0.816 - 1.0)

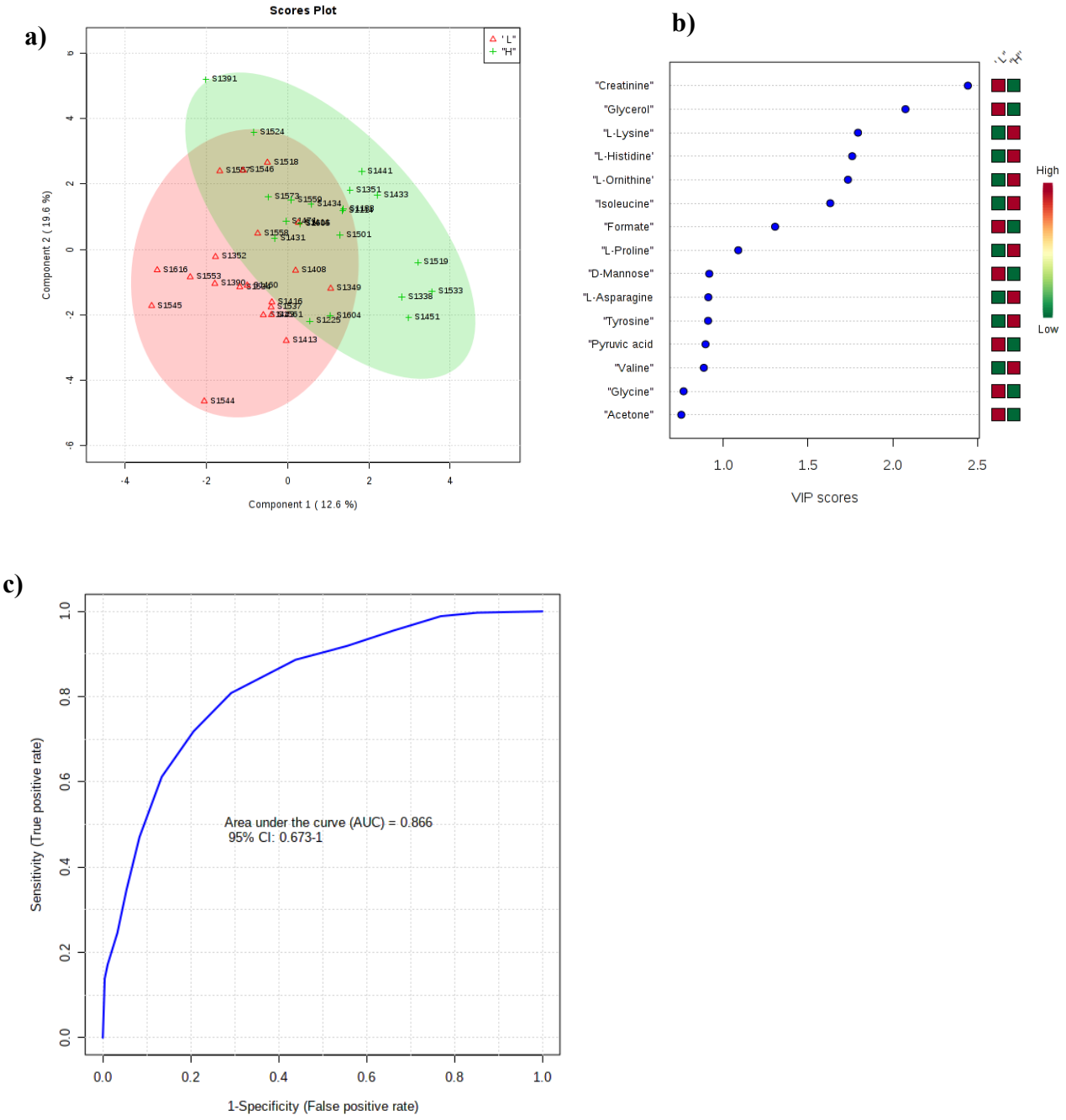

Figure 3.2 a) Score plots between first and second components of partial least square discriminant analysis (PLS-DA) on left panel and **b)** variable of importance in projection scores of PLS-DA analysis using metabolite profile of day 150 of lactation for classification of high and low RFI groups. **c)** The performance of top three serum metabolites (Creatinine, glycerol, and L-lysine) as biomarkers of RFI is presented in ROC curve. AUC = 0.866 (CI: 0.673 - 1.0)

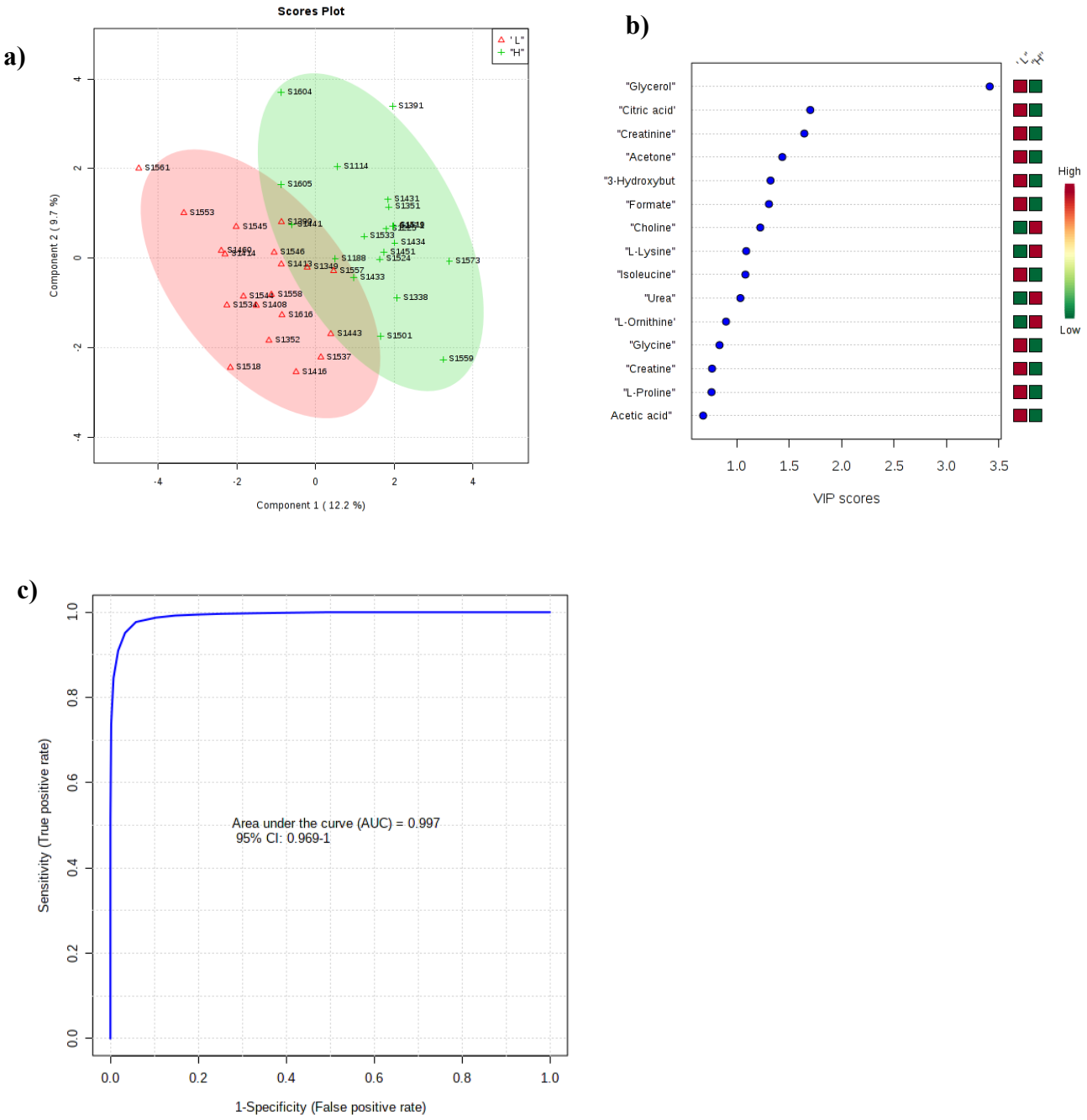

Figure 3.3 a) Score plots between first and second components of partial least square discriminant analysis (PLS-DA) on left panel and **b)** variable of importance in projection scores of PLS-DA analysis using metabolite profile of day 240 of lactation for classification of high and low RFI groups. **c)** The performance of top two serum metabolites (glycerol and creatinine) as biomarkers of RFI is presented in ROC curve. AUC = 0.997 (CI: 0.969 - 1.0).

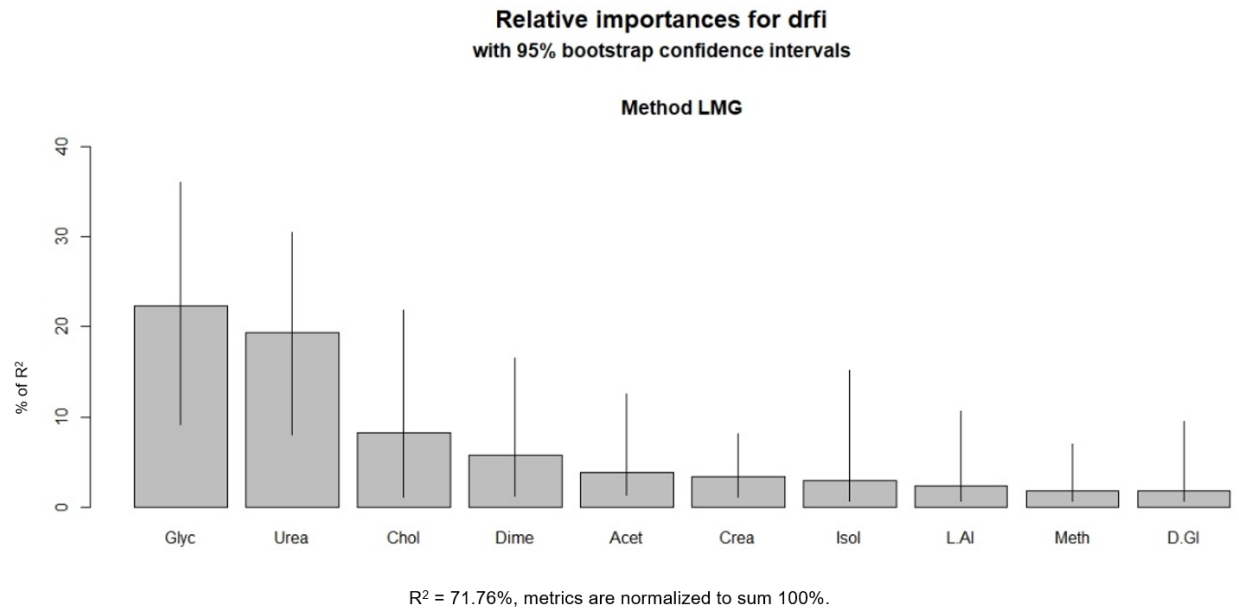

Figure 3.4 Stability of relative contribution of variables to multivariate regression models for RFI using serum metabolites from 50DIM. Values expressed as contribution to R^2 with 95% confidence intervals. Abbreviations: Gly = Glycerol; Urea = Urea; Chol = Choline; Dime = Dimethyl Sulfone; Acet = Acetone; Crea = Creatine; Isol = Isoleucine; L. Al = L-Alanine; Meth = Methionine; D. Gl = D-Glucose.

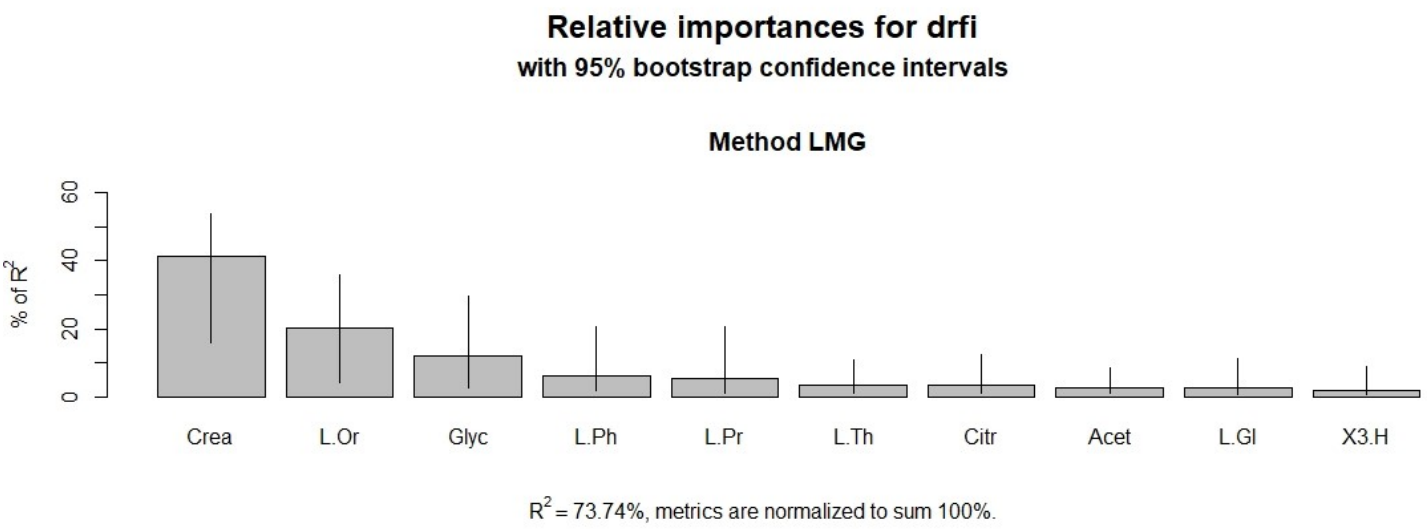

Figure 3.5 Stability of relative contribution of variables to multivariate regression models for RFI using serum metabolites from 150DIM. Values expressed as contribution to R^2 with 95% confidence intervals. Abbreviations: Crea = Creatinine; L. Or = L-Ornithine; Gly = Glycerol; L.Ph = L-Phenyl alanine; L.Th = L- Threonine; Citr = Citric acid; Acet = Acetone; L.Gl = L-Glutamine; X3.H = 3-hydroxybutyric acid.

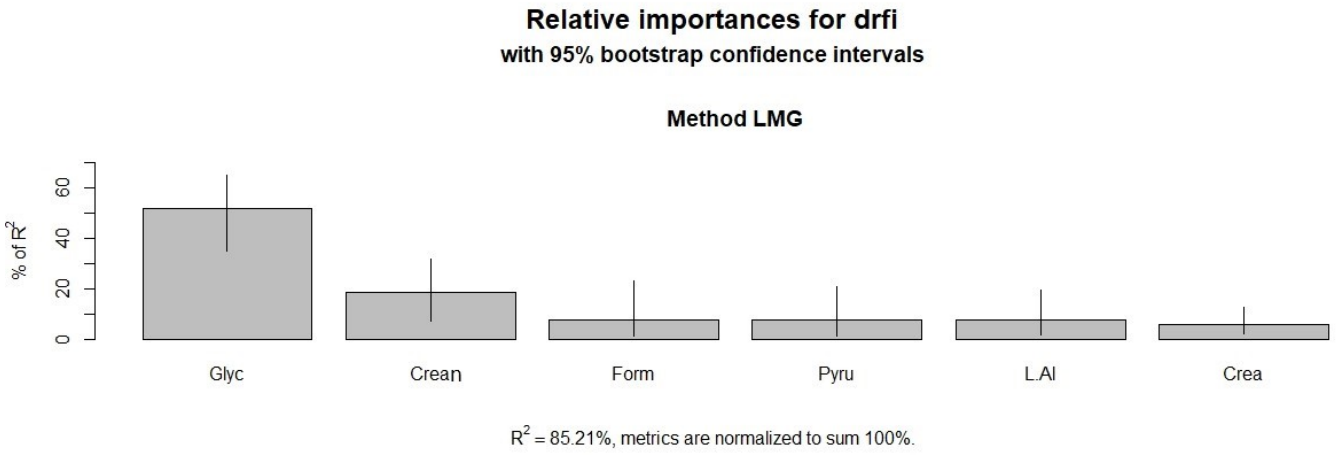

Figure 3.6 Stability of relative contribution of variables to multivariate regression models for RFI using serum metabolites from 240DIM. Values expressed as contribution to R^2 with 95% confidence intervals. Abbreviations: Gly = Glycerol; Crean = Creatinine; Form = Formate, Pyru = Pyruvic acid; L. Al = L-Alanine; Crea = Creatine.

Chapter 4. Summary and general conclusion

4.1 Summary of results

Feed cost represents over half of dairy production costs (de Haas et al. 2017). Residual feed intake has been proposed as a measure of feed efficiency (Koch, et al. 1963). Animals that consume less amount of feed for same level of production would most likely increase profitability for dairy farmers and reduce the environmental footprint of dairy production (Manafiazar et al. 2012).

Despite the efforts made to date, RFI estimation is still expensive to measure with no proposed indicator trait or classifier of animals with high or low RFI to reduce phenotyping costs. Furthermore, the biological processes underlying feed efficiency and the differences in metabolism that lead to different phenotypes is rarely studied in dairy cows. RFI of lactating dairy cows is even less studied than growing dairy calves because of the cost of recording of RFI component traits such as milk composition, milk yield, feed intake, and body weight that need to be measured along lactation.

The other aspect of RFI that few dairy studies have studied is the nonlinear changes of RFI and its component traits along lactation. These traits have nonlinear trends along lactation that can be predicted by appropriate statistical modeling. In this study, feed intake was recorded on a daily basis while milk composition and body weight were measured on a weekly and monthly basis, respectively. The daily values of the derived traits (MPER, MBW and EBW) were predicted from weekly and monthly records in order to obtain comparable daily records of all RFI component traits.

Residual feed intake (the measure of feed efficiency) and its component traits including milk production energy requirement, maintenance energy requirement and dry matter intake are

important production traits in the dairy industry (Manafiazar et al. 2013).This enabled us to calculate daily RFI values using multiple linear regression.

Animals were then grouped as high and low RFI cows (low RFI: < -0.5 SD, high RFI: > +0.5 SD) and these extreme groups were selected for metabolomics analysis. The daily changes of all derived traits for high and low RFI lactating dairy cows were plotted from day 3 to 240 of lactation. This is the first report of how RFI and its component traits change along lactation for high and low RFI lactating cows. The trend of PMBW, PEBWC, PMPER in high and low RFI groups had no significant difference because they were adjusted for estimation of RFI. But the trend of AEI and RFI were significantly different. These two plots (Figure 2.2 and 2.7) represented the gap in energy intake between high and low RFI animals that originated from their different merits for feed efficiency.

The pattern of RFI changes along lactation for high and low RFI groups described in this study (figure 2.7) agrees with a recent paper that presented the RFI plot of a herd (Li et al., 2020). This shows the differences between high and low RFI groups are relatively stable along lactation, however, further studies with a larger sample size per group of animals is needed to further validate this result. That is a contribution to the knowledge of trend of feed efficiency of divergent RFI groups in lactating dairy cows along lactation. The latest studies provide more confidence that RFI is the most relevant measure of feed efficiency as it provides a window on the underlying metabolic efficiency of dairy cows (Tempelman & Lu, 2020). This has made the investigation of the metabolic differences between high and low RFI lactating cows more interesting. Because there might be biomarkers with biological role in regulation of RFI that can classify animals by their feed efficiency. This potentially makes RFI phenotyping easier and cheaper.

In chapter 3, we ran four analyses to consider the differences in metabolite concentration between groups, the classification of groups of RFI by metabolite profile, the candidate biomarker discovery, and the prediction of RFI values using respectively univariate metabolite regression, PLS-DA classification model, ROC curve, and multiple linear regression. Similar metabolites were significant in all three types of analysis. To check validity of the result, independent analyses were done with cross validation. Although this validation was tested by accredited statistical techniques such as permutation test and bootstrap, actual validation of trained models using independent study data will show the true accuracy of our predictions.

We found 4, 4 and 8 metabolites from respectively day 50, 150 and 240 of lactation that have significantly ($P<0.05$) different concentrations between RFI groups. Glycerol and creatinine emerged as potential biomarkers for classification of RFI groups based on univariate analysis because they were significantly different in all three sampling days. Specifically, glycerol at day 50 and 240 with P-value of 10^{-6} and 10^{-12} with sample size of only 40. This indicates very little chance to see glycerol concentration be randomly different between the groups in this study. Metabolites that were significantly different were discussed and compared with literature. Pathways such as TCA cycle and urea cycle may be the origin of these differences.

Other important fixed effects such as parity and month-year showed differences for some of the metabolites. For example, urea concentration was also impacted significantly by month-year variable in the model. The effect of season was seen before in blood urea concentration of dairy cows (Alberghina et al., 2013). This was one of the differences between our study and most of other published results on RFI and metabolomics: we considered other variables that could have significant effect on metabolite concentration to avoid false interpretation about biological differences between RFI groups. This also helped us to follow the recommended method for

adjusting the metabolites for significant effects before using multiple metabolites in a statistical model (Xia et al. 2013).

Classification ability of high and low RFI animals using metabolite profile was tested using PLS-DA model. This model showed importance of metabolites when they are used together for classification. Moreover, the prediction power of models was presented by ROC curve. The AUC of ROC curves were 0.936, 0.866, and 0.997 respectively for day 50, 150 and 240 DIM. This criterion of binary classification is ideally 1.0 for a perfect classifier that does not have any false prediction. In this project, AUC values are very close to one which shows strong potential of serum metabolites for classification of RFI. This thesis has provided insights on the use of metabolites for classification of high and low RFI of lactating cows. Although many studies have tried to classify high and low RFI groups using different approaches, relatively poor success has been made.

Furthermore, in this study, we showed that a metabolomics approach has a greater potential for predicting individual RFI values in dairy cows. Our approach for regression analysis follows Karisa et al. (2014) that modeled RFI values by multiple linear regression using metabolite concentrations in beef cattle. We used this method in our study for lactating dairy cows. Linear relationship between RFI values and metabolite profile was evaluated. Feed efficiency values (RFI) for 240 days of lactation of Holsteins could be predicted with multiple regression model with high accuracy. Correlation coefficient (r) between observed and predicted RFI value (squared root of adjusted R^2) using serum metabolite concentrations from day 50, 150 and 240 of lactation were respectively 0.79, 0.80, and 0.90. This shows a strong potential for metabolites to be used as an indicator trait for RFI values in dairy cows.

Moreover, our biomarkers are metabolites that have a biological role in variation of feed efficiency, and they have potential for future studies to show biological differences between high and low RFI lactating dairy cows. The metabolites with VIP score over 1 in PLS-DA model and the candidate biomarkers of ROC curve have shown which metabolites have potential for classification of lactating cows to high and low RFI groups. The metabolites that are used as variables of linear regression model had contribution to our knowledge about RFI of lactating cows because they can be interpreted simply. The sign of the regression coefficient states how metabolites that are in the model are related to the RFI value of lactating cows. They show whether higher or lower values of a metabolite concentration is associated with more feed efficient cows. The regression coefficients of metabolites in linear models show how much RFI value (Mcal) changes with one unit change of metabolite concentration (μM). The regression coefficients with standard errors are provided in this thesis (Tables 3.5, 3.4, 3.6). Additionally, the sign of regression coefficient of metabolite represents the direction of this change. The relative contribution percentage for each metabolite with provided confidence interval indicates how much of RFI variation can be explained by variation of each metabolite in the model (Tables 3.5, 3.4, 3.6). The regression model results provided insights on the role of different metabolites on RFI values of dairy cows.

There is a potential for serum metabolites to be used as indicators of feed efficiency in dairy cows, however there are conditions to be met for practical use of metabolites. The strong relationship between metabolites and RFI must exist in large-scale dataset. This relationship must be seen in independent studies using data from different farms and labs. In other words, the strong phenotypic relationship between RFI and metabolites has to be reproducible. In addition, the indicator metabolites of RFI needs to have greater heritability than RFI. The most important

conditions to be met is less cost and ease of measurement of metabolites in comparison to conventional method of data collection of individual lactating cows along lactation.

4.2 Study limitations and future works

The number of metabolites that each metabolomic platform can identify and quantify from biological samples is limited. In this study, we analyzed only 33 metabolites that were above the limit of quantification (LOQ) from NMR platform. But that is not the whole set of metabolites that could be quantified in serum samples. There are other metabolomic platforms that can quantify more metabolites from serum samples. It is unknown if metabolite form other metabolomic platforms could provide better separation between high and low RFI lactating cows. Therefore, there could be better metabolite biomarkers associated with RFI in lactating dairy cows using other metabolomic platforms such as GC-MS, LC-MS etc. Future studies can work on the potential of metabolites as biomarkers from other metabolomic platforms for classification of high and low RFI lactating cows.

This study was limited to metabolomics and did not use other biological biomarkers from other omics such as DNA, RNA, or protein data. This limitation made the biological pathway analysis impossible because of low number of metabolites that were significantly different between high and low RFI groups. Although promising classification result with clear separation was evident in our study, the number of significantly different metabolites only from NMR platform was not enough for further investigation of differences in biological pathways between RFI groups. Therefore, combination of serum metabolomics with serum proteomics and genomics may reveal biological pathways associated with feed efficiency in dairy cows. Some proteins have enzymatic role in biological reactions. Some are necessary for mitochondrial function and energy

metabolism. There are 71 differentially expressed proteins reported in liver tissue of high and low RFI beef cattle groups (Fonseca et al., 2019). Addition of proteomics to metabolomics will make biological interpretation of differences between high and low dairy RFI cows more reliable. Because the relationship between metabolites and proteins can be investigated. Furthermore, integrated analysis of metabolomics and genes will reveal biological networks associated with feed efficiency with more reliability than combining only two omics in pathway analysis.

Future works can focus on the association of metabolite profile and RFI with similar design that we have done. This can validate the result of this study because a true validation can be done only by increase of sample size with similar experiment conditions. However statistical measures were taken before selection of adequate model in our study, future studies with similar experiment will determine reproducibility of results. The power of models for prediction using unseen data has been always influenced by sample size of population. Machine learning models such as PLS-DA and statistical models like linear regression both gain more prediction power with larger sample size. The advantage of future studies with considerably larger sample size will be narrower standard error for regression parameters.

4.3 Conclusion

Overall, this study showed how RFI and its component traits change along lactation and found a significant difference between high and low RFI cows and that this difference was stable or consistent during lactation. Candidate metabolite biomarkers that could classify high and low RFI groups of lactating cows with high accuracy were identified. We obtained AUC of 0.936, 0.866, and 0.997 in ROC curve analysis using 2-3 metabolite concentration from serum samples

at 50, 150 and 240 DIM respectively. Furthermore, the prediction of individual RFI in lactating dairy cows using serum metabolite profiles from 50, 150 and 240 DIM was possible with cross-validated R-squared of 0.52, 0.25, and 0.71 respectively. This is promising result that individual lactation RFI can be predicted with moderate accuracy 190 days before end of the trial considering the sample size of study. Further studies need to be done to validate this result with larger sample size. There is a strong potential to reduce costs and time related to regular recording of feed intake and production traits for indirect estimation of RFI. This study is a potentially important step forward for this development.

4.4 References

Alberghina, D., Piccione, G., Casella, S., Panzera, M., Morgante, M., & Gianesella, M. (2013). The effect of the season on some blood metabolites and haptoglobin in dairy cows during postpartum period. *Archives Animal Breeding*, *56*(1), 354–359. https://doi.org/10.7482/0003-9438-56-035

de Haas, Y., Pszczola, M., Soyeurt, H., Wall, E., & Lassen, J. (2017). Invited review: Phenotypes to genetically reduce greenhouse gas emissions in dairying. *Journal of Dairy Science*, *100*(2), 855–870. https://doi.org/10.3168/jds.2016-11246

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., … Fukumasu, H. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-41813-x

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., … Plastow, G. S. (2014). Plasma metabolites associated with residual feed intake and other productivity

performance traits in beef cattle. *Livestock Science*, *165*(1), 200–211. https://doi.org/10.1016/j.livsci.2014.03.002

Koch, R. M., Chambers, D., & Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *Journal of Animal Science*, *22*(2), 486–494.

Li, B., VanRaden, P. M., Guduk, E., O'Connell, J. R., Null, D. J., Connor, E. E., … Cole, J. B. (2020). Genomic prediction of residual feed intake in US Holstein dairy cattle. *Journal of Dairy Science*, *103*(3), 2477–2486. https://doi.org/10.3168/jds.2019-17332

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., Basarab, J., Li, P., & Wang, Z. (2013). Prediction of residual feed intake for first-lactation dairy cows using orthogonal polynomial random regression. *Journal of Dairy Science*, *96*(12), 7991–8001. https://doi.org/10.3168/jds.2013-6560

Manafiazar, G, McFadden, T., Goonewardene, L., Okine, E., & Wang, Z. (2012). Factors That Affect Energy Efficiency and Indirect Response of Selection for Efficiency on Related Traits. https://doi.org/http://dx.doi.org/10.5772/50776

Tempelman, R. J., & Lu, Y. (2020). Symposium review: Genetic relationships between different measures of feed efficiency and the implications for dairy cattle selection indexes. *Journal of Dairy Science*, *103*(6), 5327–5345. https://doi.org/10.3168/jds.2019-17781

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, *9*(2), 280–299. https://doi.org/10.1007/s11306-012-0482-9

Bibliography

Alberghina, D., Piccione, G., Casella, S., Panzera, M., Morgante, M., & Gianesella, M. (2013). The effect of the season on some blood metabolites and haptoglobin in dairy cows during postpartum period. *Archives Animal Breeding*, *56*(1), 354–359. https://doi.org/10.7482/0003-9438-56-035

Amaral-Phillips, D. M. (2012). *Using Peak and Summit Milk to Evaluate Your Dairy's Management Programs*. Retrieved from https://afs.ca.uky.edu/files/using_peak_and_summit_milk_to_evaluate_your_dairys_management_programs.pdf

Armbruster, D. A., & Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist. Reviews*, *29 Suppl 1*(Suppl 1), S49-52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18852857

Arndt, C., Powell, J. M., Aguerre, M. J., Crump, P. M., & Wattiaux, M. A. (2015). Feed conversion efficiency in dairy cows: Repeatability, variation in digestion and metabolism of energy and nitrogen, and ruminal methanogens. *Journal of Dairy Science*, *98*(6), 3938–3950. https://doi.org/10.3168/jds.2014-8449

Arthur, J. P. F., & Herd, R. M. (2008). Residual feed intake in beef cattle. *Revista Brasileira de Zootecnia*, *37*(SPECIALISSUE), 269–279. https://doi.org/10.1590/S1516-35982008001300031

Arthur, P. F., Archer, J. A., Johnston, D. J., Herd, R. M., Richardson, E. C., & Parnell, P. F. (2001). Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *Journal of Animal Science*, *79*(11),

2805–2811. https://doi.org/10.2527/2001.79112805x

Basarab, J. A., Beauchemin, K. A., Baron, V. S., Ominski, K. H., Guan, L. L., Miller, S. P., & Crowley, J. J. (2013). Reducing GHG emissions through genetic improvement for feed efficiency: effects on economically important traits and enteric methane production. *Animal : An International Journal of Animal Bioscience*. https://doi.org/10.1017/S1751731113000888

Belay, T. K., Dagnachew, B. S., Kowalski, Z. M., & Ådnøy, T. (2017). An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. *Journal of Dairy Science*, *100*(8), 6312–6326. https://doi.org/10.3168/jds.2016-12252

Belay, T. K., Svendsen, M., Kowalski, Z. M., & Ådnøy, T. (2017). Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows. *Journal of Dairy Science*, *100*(8), 6298–6311. https://doi.org/10.3168/jds.2016-12458

Berry, D. P., & Crowley, J. J. (2012). Residual intake and body weight gain: A new measure of efficiency in growing cattle. *Journal of Animal Science*, *90*(1), 109–115. https://doi.org/10.2527/jas.2011-4245

Berry, D. P., & Crowley, J. J. (2013). Cell biology symposium: Genetics of feed efficiency in dairy and beef cattle. *Journal of Animal Science*, *91*, 1594–1613. https://doi.org/10.2527/jas2012-5862

Bertram, H. C., Kristensen, N. B., Vestergaard, M., Jensen, S. K., Sehested, J., Nielsen, N. C., & Malmendal, A. (2009). Metabolic characterization of rumen epithelial tissue from dairy calves fed different starter diets using 1H NMR spectroscopy. *Livestock Science*, *120*(1–2),

127–134. https://doi.org/10.1016/j.livsci.2008.05.001

Betancourt, F., Lachance, H., Friedman, S., Perlee, S., & McKay, S. (2016). Exploring DNA Methylation in the Bovine 5 ’ AMPK Gene Family (p. 1). Burlington: University of Vermont. Retrieved from https://www.researchgate.net/profile/Hannah_Lachance/publication/309121378_Exploring_DNA_Methylation_in_the_Bovine_5’_AMPK_Gene_Family/links/57ffc2e308ae6fc7fc6522d0/Exploring-DNA-Methylation-in-the-Bovine-5-AMPK-Gene-Family.pdf

Bloch, H. (1946). The Nuclear induction Experiment. *Physical Review*, *70*(7).

Canadian Dairy Commission, C. D. (2016). Canadian Dairy Commission, (November). Retrieved from http://www.cdc-ccl.gc.ca/CDC/index-eng.php?id=3806

Cantalapiedra-Hijar, G., Guarnido, P., Schiphorst, A. M., Robins, R. J., Renand, G., & Ortigues-Marty, I. (2019). Natural 15N abundance in specific amino acids indicates associations between transamination rates and residual feed intake in beef cattle. *Journal of Animal Science*, *98*(6), 1–7. https://doi.org/10.1093/jas/skaa171

CCAC. (2020). Retrieved July 15, 2020, from https://www.ccac.ca/

Chen, X. B., Jayasuriya, M. C. N., & Makkar, H. P. S. (2004). Measurement and Application of Purine Derivatives:Creatinine Ratio in Spot Urine Samples of Ruminants. In *Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives* (pp. 167–179). Springer Netherlands. https://doi.org/10.1007/978-1-4020-2844-1_20

Ciborowski, M., Lipska, A., Godzien, J., Ferrarini, A., Korsak, J., Radziwon, P., … Barbas, C. (2012). Combination of LC-MS- and GC-MS-based metabolomics to study the effect of ozonated autohemotherapy on human blood. *Journal of Proteome Research*, *11*(12), 6231–6241. https://doi.org/10.1021/pr3008946

Collard, B. L., Boettcher, P. J., Dekkers, J. C. M., Petitclerc, D., & Schaeffer, L. R. (2000). Relationships between energy balance and health traits of dairy cattle in early lactation. *Journal of Dairy Science*, *83*(11), 2683–2690. https://doi.org/10.3168/jds.S0022-0302(00)75162-9

Connor, E. E. (2014). Invited review: Improving feed efficiency in dairy production: challenges and possibilities*. https://doi.org/10.1017/S1751731114002997

Connor, E. E., Hutchison, J. L., Norman, H. D., Olson, K. M., Van Tassell, C. P., Leith, J. M., & Baldwin VI, R. L. (2013). Use of residual feed intake in Holsteins during early lactation shows potential to improve feed efficiency through genetic selection. *Journal of Animal Science*, *91*(8), 3978–3988. https://doi.org/10.2527/jas.2012-5977

Consolo, N. R. B., Buarque, V. L. M., Silva, J., Poleti, M. D., Barbosa, L. C. G. S., Higuera-Padilla, A., … Silva, S. L. (2020). Muscle and liver metabolomic signatures associated with residual feed intake in Nellore cattle. *Animal Feed Science and Technology*, *8401*(20). https://doi.org/10.1016/j.anifeedsci.2020.114757

Crews, D. H. D. (2005). Genetics of efficient feed utilization and national cattle evaluation: a review. *Genetics and Molecular Research : GMR*, *4*(2), 152–165. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16110437

Crowley, J. J., McGee, M., Kenny, D. A., Crews, D. H., Evans, R. D., & Berry, D. P. (2010). Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. *Journal of Animal Science*, *88*(3), 885–894. https://doi.org/10.2527/jas.2009-1852

Dann, H. M., & Drackley, J. K. (2005). Carnitine palmitoyltransferase I in liver of periparturient dairy cows: Effects of prepartum intake, postpartum induction of ketosis, and periparturient

disorders. *Journal of Dairy Science*, *88*(11), 3851–3859. https://doi.org/10.3168/jds.S0022-0302(05)73070-8

Darrell D. Marshall and Robert Powers. (2017). Beyond the Paradigm: Combining Mass Spectrometry and Nuclear. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Dass, R., Grudziąż, K., Ishikawa, T., Nowakowski, M., Dębowska, R., & Kazimierczuk, K. (2017). Fast 2D NMR Spectroscopy for In vivo Monitoring of Bacterial Metabolism in Complex Mixtures. *Frontiers in Microbiology*, *8*(JUL), 1306. https://doi.org/10.3389/fmicb.2017.01306

Davis, S. R., Macdonald, K. A., Waghorn, G. C., & Spelman, R. J. (2014a). Residual feed intake of lactating Holstein-Friesian cows predicted from high-density genotypes and phenotyping of growing heifers. *Journal of Dairy Science*, *97*(3). https://doi.org/10.3168/jds.2013-7205

Davis, S. R., Macdonald, K. A., Waghorn, G. C., & Spelman, R. J. (2014b). Residual feed intake of lactating Holstein-Friesian cows predicted from high-density genotypes and phenotyping of growing heifers. *Journal of Dairy Science*, *97*(3), 1436–1445. https://doi.org/10.3168/jds.2013-7205

de Haas, Y., Pszczola, M., Soyeurt, H., Wall, E., & Lassen, J. (2017). Invited review: Phenotypes to genetically reduce greenhouse gas emissions in dairying. *Journal of Dairy Science*, *100*(2), 855–870. https://doi.org/10.3168/jds.2016-11246

de Ondarza, M. B., & Tricarico, J. M. (2017). Review : Advantages and limitations of dairy efficiency measures and the effects of nutrition and feeding management interventions. *The Professional Animal Scientist*, *33*(4), 393–400. https://doi.org/10.15232/pas.2017-01624

Dechow, C. D., Baumrucker, C. R., Bruckmaier, R. M., & Blum, J. W. (2017). Blood plasma

traits associated with genetic merit for feed utilization in Holstein cows. *Journal of Dairy Science*, *100*(10), 8232–8238. https://doi.org/10.3168/jds.2016-12502

Derno, M., Nürnberg, G., & Kuhla, B. (2019). Characterizing the metabotype and its persistency in lactating Holstein cows: An approach toward metabolic efficiency measures. *Journal of Dairy Science*, *102*(7), 6559–6570. https://doi.org/10.3168/jds.2019-16274

Donkin, S. S. (2008). Glycerol from biodiesel production: The new corn for dairy cattle. *Revista Brasileira de Zootecnia*, *37*(SPECIALISSUE), 280–286. https://doi.org/10.1590/S1516-35982008001300032

Dórea, J. R. R., Danés, M. A. C., Zanton, G. I., & Armentano, L. E. (2017). Urinary purine derivatives as a tool to estimate dry matter intake in cattle: A meta-analysis. *Journal of Dairy Science*, *100*(11), 8977–8994. https://doi.org/10.3168/jds.2017-12908

Dunn, W. B., & Ellis, D. I. (2005). Metabolomics: Current analytical platforms and methodologies. *TrAC - Trends in Analytical Chemistry*, *24*(4), 285–294. https://doi.org/10.1016/j.trac.2004.11.021

Eastridge, M. L. (2002). Energy in the New Dairy NRC (p. 11). Feedstuffs. Retrieved from https://dairy.osu.edu/sites/dairy/files/imce/PDF/Feed_PDF/Energy New Dairy NRC.pdf

Elolimy, A. A., Arroyo, J. M., Batistel, F., Iakiviak, M. A., & Loor, J. J. (2018). Association of residual feed intake with abundance of ruminal bacteria and biopolymer hydrolyzing enzyme activities during the peripartal period and early lactation in Holstein dairy cows, 1–10. https://doi.org/10.1186/s40104-018-0258-9

Emwas, A. H., Luchinat, C., Turano, P., Tenori, L., Roy, R., Salek, R. M., … Wishart, D. S. (2015, November 21). Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review.

Metabolomics. Springer New York LLC. https://doi.org/10.1007/s11306-014-0746-7

Emwas, A. H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods in Molecular Biology*, *1277*, 161–193. https://doi.org/10.1007/978-1-4939-2377-9_13

Fiehn, O., Wohlgemuth, G., & Scholz, M. (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. In *Lecture Notes in Bioinformatics (Subseries of Lecture Notes in Computer Science)* (Vol. 3615, pp. 224–239). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11530084_18

Fisher, L. J., Erfle, J. D., Lodge, G. A., & Sauer, F. D. (1973). Effects of propylene glycol or glycerol supplementation on feed intake, milk yield and composition, and incidence of ketosis. *Canadian Journal of Animal Science*, *53*(2), 289–296. https://doi.org/10.4141/cjas73-045

Flay, H. E., Kuhn-Sherlock, B., Macdonald, K. A., Camara, M., Lopez-Villalobos, N., Donaghy, D. J., & Roche, J. R. (2019). Hot topic: Selecting cattle for low residual feed intake did not affect daily methane production but increased methane yield. *Journal of Dairy Science*, *102*(3), 2708–2713. https://doi.org/10.3168/jds.2018-15234

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., … Fukumasu, H. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-41813-x

Foroutan, A., Guo, A. C., Vazquez-Fresno, R., Lipfert, M., Zhang, L., Zheng, J., … Wishart, D. S. (2019). Chemical Composition of Commercial Cow's Milk. *Journal of Agricultural and Food Chemistry*, *67*(17), 4897–4914. https://doi.org/10.1021/acs.jafc.9b00204

French, P. (2017). *Applied Animal Nutrition Feedstuffs* (3rd Editio). Upper Saddle River, NJ, United States: Pearson Education (US). Retrieved from https://courses.ecampus.oregonstate.edu/ans312/seven/cows_6_trans.htm

Gaines, W. L. (1946). Live Weight Versus Metabolic Body Size in Dairy Cows and Goats. *Journal of Dairy Science*, *29*(5), 259–272. https://doi.org/10.3168/jds.S0022-0302(46)92477-0

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS ONE*, *12*(5), 1–26. https://doi.org/10.1371/journal.pone.0177675

Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., & Deigner, H.-P. (2010). Targeted Metabolomics for Biomarker Discovery. *Angewandte Chemie International Edition*, *49*(32), 5426–5445. https://doi.org/10.1002/anie.200905579

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, *17*(1), 1–27. https://doi.org/10.18637/jss.v017.i01

Guo, J., Zhang, M., Elmore, C. S., & Vishwanathan, K. (2013). An integrated strategy for in vivo metabolite profiling using high-resolution mass spectrometry based data processing techniques. *Analytica Chimica Acta*, *780*, 55–64. https://doi.org/10.1016/j.aca.2013.04.012

Hailemariam, D., Mandal, R., Saleem, F., Dunn, S. M., Wishart, D. S., & Ametaj, B. N. (2014). Identification of predictive biomarkers of disease state in transition dairy cows. *Journal of Dairy Science*, *97*(5), 2680–2693. https://doi.org/10.3168/jds.2013-6803

Hardie, L C, VandeHaar, M. J., Tempelman, R. J., Weigel, K. A., Armentano, L. E., Wiggans, G. R., … Spurlock, D. M. (2017). The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. *Journal of Dairy Science*, *100*(11), 9061–9075.

https://doi.org/10.3168/jds.2017-12604

Hardie, L. C., Beitz, D. C., Fernando, R. L., & Krizan, Z. (2016). The genetic basis and improvement of feed efficiency in lactating Holstein dairy cattle. Retrieved from http://search.proquest.com.ezproxy.library.uq.edu.au/docview/1872300140/

Hassoun, P. (2019). Residual feed intake in dairy ewes : an evidence of intraflock variability. *BioRxiv*, (August), 723809. https://doi.org/10.1101/723809

Hayhurst, C., Flint, A. P. F., Løvendahl, P., Woolliams, J. A., & Royal, M. D. (2009). Genetic variation of metabolite and hormone concentration in UK Holstein-Friesian calves and the genetic relationship with economically important traits. *Journal of Dairy Science*, *92*(8), 4001–4007. https://doi.org/10.3168/jds.2008-1130

Hayhurst, C., Sørensen, M. K., Royal, M. D., & Løvendahl, P. (2007). Metabolic regulation in Danish bull calves and the relationship to the fertility of their female offspring. *Journal of Dairy Science*, *90*(8), 3909–3916. https://doi.org/10.3168/jds.2006-731

Hemme, T., & Otte, J. (2010). Pro-Poor Livestock Policy Initiative Status and Prospects for Smallholder Milk Production A Global Perspective. *Food and Agriculture Organization of the United Nations*, 1–12.

Herd, R. M., & Arthur, P. F. (2009). Physiological basis for residual feed intake. *Journal of Animal Science*, *87*(14 Suppl), 64–71. https://doi.org/10.2527/jas.2008-1345

Hungate, R. E., Smith, W., Bauchop, T., Yu, I., & Rabinowitz, J. C. (1970). Formate as an intermediate in the bovine rumen fermentation. *Journal of Bacteriology*, *102*(2), 389–397. https://doi.org/10.1128/jb.102.2.389-397.1970

Jones, E. J. (2009). Using Heritability for Genetic Improvement. Retrieved October 10, 2020, from https://www.pubs.ext.vt.edu/404/404-084/404-084.html

Jorge-Smeding, E., Renand, G., Centeno, D., Pétéra, M., Durand, S., Polakof, S., & Cantalapiedra-Hijar, G. (2019). Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. In *International Symposium on Energy and Protein Metabolism and Nutrition (ISEP)* (pp. 231–232). https://doi.org/10.3920/978-90-8686-891-9_50

Kadarmideen, H. N. (2008a). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Kadarmideen, H. N. (2008b). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Kalapos, M. P. (2003, May 2). On the mammalian acetone metabolism: From chemistry to clinical implications. *Biochimica et Biophysica Acta - General Subjects*. Elsevier. https://doi.org/10.1016/S0304-4165(03)00051-5

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., … Plastow, G. S. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livestock Science*, *165*(1), 200–211. https://doi.org/10.1016/j.livsci.2014.03.002

Karisa, B., Moore, S., & Plastow, G. (2013). Complicated Relationships: A Review of Biological Interaction Networks and Pathways in Animal Science. *Springer Science Reviews*, *1*(1–2), 73–83. https://doi.org/10.1007/s40362-013-0005-8

Karisa, B., Moore, S., & Plastow, G. (2014). Analysis of biological networks and biological pathways associated with residual feed intake in beef cattle. *Animal Science Journal*, *85*(4),

374–387. https://doi.org/10.1111/asj.12159

Karisa, Brian K. (2013). Candidate Genes , Metabolites and Biological Pathways Associated With Residual Feed Intake and Carcass. (Publication No. NS27583.) [PhD thesis, University of Alabama]. ProQuest Dissertations & ThesesGlobal

Kennedy, B. W., van der Werf, J. H., & Meuwissen, T. H. (1993). Genetic and statistical properties of residual feed intake. *Journal of Animal Science*, *71*(12), 3239–3250. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8294275

Kennedy, P. M., & Milligan, L. P. (1980). the Degradation and Utilization of Endogenous Urea in the Gastrointestinal Tract of Ruminants: a Review. *Canadian Journal of Animal Science*, *60*(2), 205–221. https://doi.org/10.4141/cjas80-030

Khansefid, M., Millen, C. A., Chen, Y., Pryce, J. E., Chamberlain, A. J., Vander Jagt, C. J., … Goddard, M. E. (2017). Gene expression analysis of blood, liver, and muscle in cattle divergentlya selected for high and low residual feed intake. *Journal of Animal Science*, *95*(11), 4764–4775. https://doi.org/10.2527/jas2016.1320

Kholif, A. E. (2019). Glycerol use in dairy diets: A systemic review. *Animal Nutrition*, *5*(3), 209–216. https://doi.org/10.1016/j.aninu.2019.06.002

Koch, R. M., Chambers, D., & Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *Journal of Animal Science*, *22*(2), 486–494.

Koeck, A., Jamrozik, J., Kistemaker, G. J., Schenkel, F. S., Moore, R. K., Lefebvre, D. M., … Miglior, F. (2016). Genetic and phenotypic associations of milk β-hydroxybutyrate with ketosis in Canadian Holsteins. *Canadian Journal of Animal Science*, *96*(3), 302–305. https://doi.org/10.1139/cjas-2015-0206

Koeslag, J. H., Noakes, T. D., & Sloan, A. W. (1980). Post-exercise ketosis. *The Journal of*

Physiology, *301*(1), 79–90. https://doi.org/10.1113/jphysiol.1980.sp013190

Korver, S., van Eekelen, E. A. M., Vos, H., Nieuwhof, G. J., & van Arendonk, J. A. M. (1991). Genetic parameters for feed intake and feed efficiency in growing dairy heifers. *Livestock Production Science*, *29*(1), 49–59. https://doi.org/10.1016/0301-6226(91)90119-B

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1). https://doi.org/10.18637/jss.v069.i01

Li, B., Berglund, B., Fikse, W. F., Lassen, J., Lidauer, M. H., Mäntysaari, P., & Løvendahl, P. (2017). Neglect of lactation stage leads to naive assessment of residual feed intake in dairy cattle. *Journal of Dairy Science*, *100*(11), 9076–9084. https://doi.org/10.3168/jds.2017-12775

Li, B., VanRaden, P. M., Guduk, E., O'Connell, J. R., Null, D. J., Connor, E. E., … Cole, J. B. (2020). Genomic prediction of residual feed intake in US Holstein dairy cattle. *Journal of Dairy Science*, *103*(3), 2477–2486. https://doi.org/10.3168/jds.2019-17332

Linn, J. (2003). Energy in the 2001 Dairy NRC : Understanding the System, (Figure 1), 102–110.

Maher, A. D., Hayes, B., Cocks, B., Marett, L., Wales, W. J., & Rochfort, S. J. (2013). Latent biochemical relationships in the blood-milk metabolic axis of dairy cows revealed by statistical integration of 1H NMR spectroscopic data. *Journal of Proteome Research*, *12*(3), 1428–1435. https://doi.org/10.1021/pr301056q

Manafiazar, G., Baron, V. S., McKeown, L., Block, H., Ominski, K., Plastow, G., & Basarab, J. A. (2020). Methane and carbon dioxide emissions from yearling beef heifers and mature cows classified for residual feed intake under drylot conditions. *Canadian Journal of Animal Science*, *100*(3), 522–535. https://doi.org/10.1139/cjas-2019-0032

Manafiazar, G., Goonewardene, L., Miglior, F., Crews, D. H., Basarab, J. A., Okine, E., &

Wang, Z. (2015a). Genetic and phenotypic correlations among feed efficiency, production and selected conformation traits in dairy cows. *Animal*, *10*(3). https://doi.org/10.1017/S1751731115002281

Manafiazar, G., Goonewardene, L., Miglior, F., Crews, D. H., Basarab, J. A., Okine, E., & Wang, Z. (2015b). Genetic and phenotypic correlations among feed efficiency, production and selected conformation traits in dairy cows. *Animal*, *10*(3), 381–389. https://doi.org/10.1017/S1751731115002281

Manafiazar, G., McFadden, T., Goonewardene, L., Okine, E., Basarab, J., Li, P., & Wang, Z. (2013). Prediction of residual feed intake for first-lactation dairy cows using orthogonal polynomial random regression. *Journal of Dairy Science*, *96*(12), 7991–8001. https://doi.org/10.3168/jds.2013-6560

Manafiazar, G. (2013). Investigation of Residual Feed Intake in Dairy. (Publication No. NS27639.) [PhD thesis, University of Alabama]. ProQuest Dissertations & ThesesGlobal.

Manafiazar, G, McFadden, T., Goonewardene, L., Okine, E., & Wang, Z. (2012). Factors That Affect Energy Efficiency and Indirect Response of Selection for Efficiency on Related Traits. https://doi.org/http://dx.doi.org/10.5772/50776

Martin, O., & Sauvant, D. (2007). Dynamic model of the lactating dairy cow metabolism. *Animal*, *1*(8), 1143–1166. https://doi.org/10.1017/S1751731107000377

Kuhn, M. (2005). Building Predictive Models in R Using the caret Package. *Journal Of Statistical Software*, *14*(September), 981029. https://doi.org/10.18637/jss.v081.b02

McDonald, P., Edwards, R. A., Greenhalgh, J. F. D., & Morgan, C. A. (2002). Animal Nutrition. (7th Editio). Pearson Canada.

McParland, S., Lewis, E., Kennedy, E., Moore, S. G., McCarthy, B., O'Donovan, M., … Berry,

D. P. (2014). Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. *Journal of Dairy Science*, *97*(9), 5863–5871. https://doi.org/10.3168/jds.2014-8214

Mellado, M., Antonio-Chirino, E., Meza-Herrera, C., Veliz, F. G., Arevalo, J. R., Mellado, J., & de Santiago, A. (2011). Effect of lactation number, year, and season of initiation of lactation on milk yield of cows hormonally induced into lactation and treated with recombinant bovine somatotropin. *Journal of Dairy Science*, *94*(9), 4524–4530. https://doi.org/10.3168/jds.2011-4152

Miglior, F., Fleming, A., Malchiodi, F., Brito, L. F., Martin, P., & Baes, C. F. (2017). A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *Journal of Dairy Science*, *100*(12), 10251–10271. https://doi.org/10.3168/jds.2017-12968

Moe, P. W. (1981). Energy Metabolism of Dairy Cattle. *Journal of Dairy Science*, *64*(6), 1120–1139. https://doi.org/10.3168/jds.S0022-0302(81)82692-6

Montanholi, Y. R., Haas, L. S., Swanson, K. C., Coomber, B. L., Yamashiro, S., & Miller, S. P. (2017). Liver morphometrics and metabolic blood profile across divergent phenotypes for feed efficiency in the bovine. *Acta Veterinaria Scandinavica*, *59*(1), 1–11. https://doi.org/10.1186/s13028-017-0292-1

Mulligan, F., O'Grady, L., Rice, D., & Doherty, M. (2006). Production diseases of the transition cow: Body condition score and energy balance. *Irish Veterinary Journal*, *59*(9), 505–510.

Neves, R. C., Leno, B. M., Curler, M. D., Thomas, M. J., Overton, T. R., & McArt, J. A. A. (2018). Association of immediate postpartum plasma calcium concentration with early-lactation clinical diseases, culling, reproduction, and milk production in Holstein cows.

Journal of Dairy Science, *101*(1), 547–555. https://doi.org/10.3168/jds.2017-13313

Newman, J. C., & Verdin, E. (2017). β-Hydroxybutyrate: A Signaling Metabolite. *Annual Review of Nutrition*, *37*, 51–76. https://doi.org/10.1146/annurev-nutr-071816-064916

Nielsen, J., & Oliver, S. (2005). The next wave in metabolome analysis. *Trends in Biotechnology*, *23*(11), 544–546. https://doi.org/10.1016/j.tibtech.2005.08.005

Nkrumah, J. D. Okine, E. K., Mathison, G. W., Schmid, K., Li, C., Basarab, J. A., … Moore, S. S. (2006). Relationships of feedlot feed efficiency, performance, and feeding behaviour with metabolic rate, methane produciton, and energy partitioning in beef cattle. Journal of Animal Science, 84(1), 145–153. Retrieved from https://doi.org/10.2527/2006.841145x

O'Sullivan, M., Butler, S. T., Pierce, K. M., Crowe, M. A., O'Sullivan, K., Fitzgerald, R., & Buckley, F. (2020). Reproductive efficiency and survival of Holstein-Friesian cows of divergent Economic Breeding Index, evaluated under seasonal calving pasture-based management. *Journal of Dairy Science*, *103*(2), 1685–1700. https://doi.org/10.3168/jds.2019-17374

Osman, M. A., Allen, P. S., Mehyar, N. A., Bobe, G., Coetzee, J. F., Koehler, K. J., & Beitz, D. C. (2008). Acute metabolic responses of postpartal dairy cows to subcutaneous glucagon injections, oral glycerol, or both. *Journal of Dairy Science*, *91*(9), 3311–3322. https://doi.org/10.3168/jds.2008-0997

Porcu, C., Manca, C., Cabiddu, A., Dattena, M., Gallus, M., Pasciu, V., … Molle, G. (2018). Effects of short-term administration of a glucogenic mixture at mating on feed intake, metabolism, milk yield and reproductive performance of lactating dairy ewes. *Animal Feed Science and Technology*, *243*, 10–21. https://doi.org/10.1016/j.anifeedsci.2018.06.012

Pryce, J. E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W. J., Coffey, M. P., Hayes, B. J., &

Goddard, M. E. (2015). Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. *Journal of Dairy Science*, *98*(10), 7340–7350. https://doi.org/10.3168/jds.2015-9621

Pryce, J. E., Wales, W. J., De Haas, Y., Veerkamp, R. F., & Hayes, B. J. (2014). Genomic selection for feed efficiency in dairy cattle. *Animal*, *8*(1), 1–10. https://doi.org/10.1017/S1751731113001687

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., … Wishart, D. S. (2011). The human serum metabolome. *PLoS ONE*, *6*(2). https://doi.org/10.1371/journal.pone.0016957

Pulido, R. G., & Briones, M. (2004). Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives. *Estimation of Microbial Protein Supply in Ruminants Using Urinary Purine Derivatives*, (June 2018). https://doi.org/10.1007/978-1-4020-2844-1

R-Core-Team. (2019). R: A language and environment for statistical computing. Retrieved April 22, 2020, from https://www.r-project.org/

Ravanbakhsh, S., Liu, P., Bjordahl, T. C., Mandal, R., Grant, J. R., Wilson, M., … Wishart, D. S. (2015). Accurate, fully-automated NMR spectral profiling for metabolomics. *PLoS ONE*, *10*(5), 1–15. https://doi.org/10.1371/journal.pone.0124219

Richardson, E. C., & Herd, R. M. B. (2004). Biological basis for variation in residual feed intake in beef cattle . 2 . Synthesis of results following divergent selection Cooperative Research Centre for Cattle and Beef Quality . *Australian Journal of Experimental Agriculture*, *44*, 431–440. https://doi.org/10.1071/EA02221

Rius, A. G., Kittelmann, S., Macdonald, K. A., Waghorn, G. C., Janssen, P. H., & Sikkema, E. (2012). Nitrogen metabolism and rumen microbial enumeration in lactating cows with

divergent residual feed intake fed high-digestibility pasture. *Journal of Dairy Science*, *95*(9), 5024–5034. https://doi.org/10.3168/jds.2012-5392

Saleem, A. M., Zanouny, A. I., & Singar, A. M. (2018). Effect of glycerol supplementation during early lactation on milk yield, milk composition, nutrient digestibility and blood metabolites of dairy buffaloes. *Animal*, *12*(4), 757–763. https://doi.org/10.1017/S175173111700180X

Saleem, F., Bouatra, S., Guo, A. C., Psychogios, N., Mandal, R., Dunn, S. M., … Wishart, D. S. (2013). The Bovine Ruminal Fluid Metabolome. *Metabolomics*, *9*(2), 360–378. https://doi.org/10.1007/s11306-012-0458-9

Salleh, M. S., Mazzoni, G., Höglund, J. K., Olijhoek, D. W., Lund, P., Løvendahl, P., & Kadarmideen, H. N. (2017). RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high- and low-residual feed intake in Nordic dairy cattle. *BMC Genomics*, *18*(1), 1–17. https://doi.org/10.1186/s12864-017-3622-9

Saude, E. J., Slupsky, C. M., & Sykes, B. D. (2006). Optimization of NMR analysis of biological fluids for quantitative accuracy. *Metabolomics*, *2*(3), 113–123. https://doi.org/10.1007/s11306-006-0023-5

Schaefer, A. L., Ominski, K., Thompson, S., Crow, G., Bench, C., Colyn, J., … von Gaza, H. (2018). Energy utilization in cattle with steady state and non-steady state methods: the importance of thermal neutrality. *Heliyon*, *4*(10). https://doi.org/10.1016/j.heliyon.2018.e00843

Serkova, N. J., Standiford, T. J., & Stringer, K. A. (2011). The emerging field of quantitative blood metabolomics for biomarker discovery in critical illnesses. *American Journal of*

Respiratory and Critical Care Medicine, *184*(6), 647–655. https://doi.org/10.1164/rccm.201103-0474CI

Shetty, N., Løvendahl, P., Lund, M. S., & Buitenhuis, A. J. (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. https://doi.org/10.3168/jds.2016-11609

Solem, C., Koebmann, B., & Jensen, P. R. (2008). Genetical systems biology in livestock: application to gonadotrophin releasing hormone and reproduction. *IET Systems Biology*, *2*(2), 64–72. https://doi.org/10.1049/iet-syb

Sun, H. Z., Wang, D. M., Wang, B., Wang, J. K., Liu, H. Y., Guan, L. L., & Liu, J. X. (2015). Metabolomics of four biofluids from dairy cows: Potential biomarkers for milk production and quality. *Journal of Proteome Research*, *14*(2), 1287–1298. https://doi.org/10.1021/pr501305g

Sundekilde, U. K. (2012). Milk metabolite variability and heritability and their association with technological properties of bovine milk elucidated by NMR-based metabonomics, (July), 228.

Szymańska, E., Saccenti, E., Smilde, A. K., & Westerhuis, J. A. (2012). Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*, *8*(S1), 3–16. https://doi.org/10.1007/s11306-011-0330-3

Tempelman, R. J., & Lu, Y. (2020). Symposium review: Genetic relationships between different measures of feed efficiency and the implications for dairy cattle selection indexes. *Journal of Dairy Science*, *103*(6), 5327–5345. https://doi.org/10.3168/jds.2019-17781

Thévenot, E. A., Roux, A., Xu, Y., Ezan, E., & Junot, C. (2015). Analysis of the Human Adult

Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *Journal of Proteome Research*, *14*(8), 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

Thoh, D., Pakdeechanuan, P., & Chanjula, P. (2017). Effect of supplementary glycerin on milk composition and heat stability in dairy goats. *Asian-Australasian Journal of Animal Sciences*, *30*(12), 1711–1717. https://doi.org/10.5713/ajas.17.0066

Vadla, N. C., Davalagar, V. D., & Sripadi, P. (2013, June). Detection and characterization of N-alkyl diethanolamines and N-2-alkoxyethyl diethanolamines in milk by electrospray ionization mass spectrometry. *Metabolomics*. https://doi.org/10.1007/s11306-012-0492-7

van Knegsel, A. T. M., van der Drift, S. G. A., Horneman, M., de Roos, A. P. W., Kemp, B., & Graat, E. A. M. (2010). Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. *Journal of Dairy Science*, *93*(7), 3065–3069. https://doi.org/10.3168/jds.2009-2847

Vandehaar, M. J. (1998). Efficiency of Nutrient Use and Relationship to Profitability on Dairy Farms. *Journal of Dairy Science*, *81*(1), 272–282. https://doi.org/10.3168/jds.S0022-0302(98)75576-6

VandeHaar, M. J., Armentano, L. E., Weigel, K., Spurlock, D. M., Tempelman, R. J., & 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. https://doi.org/10.3168/jds.2015-10352

VandeHaar, M. J., & St-Pierre, N. (2006). Major Advances in Nutrition: Relevance to the Sustainability of the Dairy Industry. *Journal of Dairy Science*, *89*(4), 1280–1291.

https://doi.org/10.3168/jds.S0022-0302(06)72196-8

Veerkamp, R. F., & Emmans, G. C. (1995). Sources of genetic variation in energetic efficiency of dairy cows. *Livestock Production Science*, *44*(2), 87–97. https://doi.org/10.1016/0301-6226(95)00065-0

Wallén, S. E., Prestløkken, E., Meuwissen, T. H. E., McParland, S., & Berry, D. P. (2018). Milk mid-infrared spectral data as a tool to predict feed intake in lactating Norwegian Red dairy cows. *Journal of Dairy Science*, *101*(7), 6232–6243. https://doi.org/10.3168/jds.2017-13874

Wang, Xiao, & Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites*, *9*(7), 1–16. https://doi.org/10.3390/metabo9070151

Wang, Xiaohang, Han, W., Yang, J., Westaway, D., & Li, L. (2019). Development of chemical isotope labeling LC-MS for tissue metabolomics and its application for brain and liver metabolome profiling in Alzheimer's disease mouse model. *Analytica Chimica Acta*, *1050*, 95–104. https://doi.org/10.1016/j.aca.2018.10.060

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., … Li, C. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: I: feed efficiency and component traits. *BMC Genomics*, *21*(1), 1–22. https://doi.org/10.1186/s12864-019-6362-1

Watford, M. (2003). The urea cycle: Teaching intermediary metabolism in a physiological setting. *Biochemistry and Molecular Biology Education*, *31*(5), 289–297. https://doi.org/10.1002/bmb.2003.494031050249

Winter, G., & Krömer, J. O. (2013). Fluxomics - connecting 'omics analysis and phenotypes. *Environmental Microbiology*, *15*(7), 1901–1916. https://doi.org/10.1111/1462-2920.12064

Wishart, D. S. (2005). Metabolomics: The principles and potential applications to transplantation. *American Journal of Transplantation*, *5*(12), 2814–2820. https://doi.org/10.1111/j.1600-6143.2005.01119.x

Wishart, David S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., … Scalbert, A. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research*, *41*(D1), 801–807. https://doi.org/10.1093/nar/gks1065

Wittenburg, D., Melzer, N., Willmitzer, L., Lisec, J., Kesting, U., Reinsch, N., & Repsilber, D. (2013). Milk metabolites and their genetic variability. *Journal of Dairy Science*, *96*(4), 2557–2569. https://doi.org/10.3168/jds.2012-5635

Wood, E. (2000). Lehninger's principles of biochemistry. *Biochemistry and Molecular Biology Education*, *28*(5), 286–287. https://doi.org/10.1016/s1470-8175(00)00020-5

Worley, B., & Powers, R. (2013). Multivariate Analysis in Metabolomics. *Current Metabolomics*, *1*(1), 92–107. https://doi.org/10.2174/2213235x11301010092

Xi, Y. M., Wu, F., Zhao, D. Q., Yang, Z., Li, L., Han, Z. Y., & Wang, G. L. (2016). Biological mechanisms related to differences in residual feed intake in dairy cows. *Animal*, *10*(8), 1311–1318. https://doi.org/10.1017/S1751731116000343

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013a). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, *9*(2), 280–299. https://doi.org/10.1007/s11306-012-0482-9

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013b). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics*, *9*(2), 280–299. https://doi.org/10.1007/s11306-012-0482-9

Xia, J., Sinelnikov, I. V., Han, B., & Wishart, D. S. (2015). MetaboAnalyst 3.0-making

metabolomics more meaningful. *Nucleic Acids Research*, *43*(W1), W251–W257. https://doi.org/10.1093/nar/gkv380

Yao, C., Spurlock, D. M., Armentano, L. E., Page, C. D., VandeHaar, M. J., Bickhart, D. M., & Weigel, K. A. (2013). Random Forests approach for identifying additive and epistatic single nucleotide polymorphisms associated with residual feed intake in dairy cattle. *Journal of Dairy Science*, *96*(10), 6716–6729. https://doi.org/10.3168/jds.2012-6237

Yerushalmy, J. (1947). Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Reports*, *62*(40), 1432–1449. https://doi.org/10.2307/4586294

Yu, P., Meier, J. A., Christensen, D. A., Rossnagel, B. G., & McKinnon, J. J. (2003). Using the NRC-2001 model and the DVE/OEB system to evaluate nutritive values of Harrington (malting-type) and Valier (feed-type) barley for ruminants. *Animal Feed Science and Technology*, *107*(1–4), 45–60. https://doi.org/10.1016/S0377-8401(03)00062-2

Yu, Z., Kastenmüller, G., He, Y., Belcredi, P., Möller, G., Prehn, C., … Wang-Sattler, R. (2011). Differences between human plasma and serum metabolite profiles. *PLoS ONE*, *6*(7), 1–6. https://doi.org/10.1371/journal.pone.0021230

Zampieri, G., Vijayakumar, S., Yaneske, E., & Angione, C. (2019). Machine and deep learning meet genome-scale metabolic modeling. *PLoS Computational Biology*, *15*(7), 1–24. https://doi.org/10.1371/journal.pcbi.1007084

Zanton, G. I. (2016). Analysis of production responses to changing crude protein levels in lactating dairy cow diets when evaluated in continuous or change-over experimental designs. *Journal of Dairy Science*, *99*(6), 4398–4410. https://doi.org/10.3168/jds.2015-10438