Application of Multi-Omics Tools to Study the Genetic Background of Economically Relevant Traits in Commercial Beef Production

by

Jiyuan Li

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

Doctor of Philosophy

in

Animal Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

The sustainability and profitability of beef cattle production are largely associated with feed efficiency, carcass merit, and resistance to infectious diseases. These traits are difficult or expensive to measure on individual animals, which makes them suitable for genomic application. Currently accuracies of genomic prediction (a method that could predict genetic merit of animals based on DNA markers) for these traits are relatively low, hindering their uptake in beef cattle. The transcriptome and metabolome are intermediate, molecular phenotypes lying between genomic and phenotypic levels, which could be used to provide a better understanding of the genetic background of traits. They may therefore contribute to the development of more effective genomic selection strategies to further enhance genomic selection in beef cattle. In this thesis, integrative analyses of multi-omics data were applied to give insights into these questions.

In the first study the genetic architecture of blood metabolites was evaluated. Eleven metabolites with heritability estimates ranging from 0.09 ± 0.15 to 0.36 ± 0.15 were found. Several regions were identified that explained a small proportion of heritable genetic variation (0.62% - 4.21%). These results provided evidence for genetic variation of blood metabolites in beef cattle, and baseline information for research into the utilization of plasma metabolites for genetic improvement of beef cattle. Secondly, multiple metabolites were found to be associated with feed efficiency and carcass merit traits. Combining the results of metabolome-genome-wide association analysis identified many significant SNPs and candidate genes associated with these traits. Functional SNPs and genes are recommended to be included in SNP panels to improve the accuracy of genomic evaluation and prediction. Additionally, candidate genes were subjected to functional enrichment analyses. Several significant biological processes and networks such as lipid metabolism were identified to be associated with these important traits, which could assist

preselection or prioritizing of SNPs used in genomic prediction models. In general, the integrative analysis of genomic and metabolomic data sheds light on how genes affect phenotypes by modifying the synthesis or degradation of related metabolites and improves understanding of genetic influence on phenotypes.

Lastly, transcriptomic and genotypic data were analyzed to study the genetics of bovine respiratory disease (BRD), the most common and costly infectious disease of beef cattle in North America. BRD susceptibility showed a moderate heritability (0.43 ± 0.51) in feedlot cattle. Two significant SNPs were identified to be associated with BRD susceptibility and 101 genes which were mainly involved in inflammatory response were differentially expressed (DE) in BRD and non-BRD animals. A total of 420 cis-expression quantitative loci (cis-eQTLs) and 144 trans-eQTLs were associated with the expression of the DE genes. Investigations into the relationship between different omics levels, revealed effect of genotype on gene expression and their roles in the host immune responses and disease susceptibility. Transcriptomic biomarkers with high accuracy and reliability to predict BRD status were identified which could be used to help diagnose BRD in feedlots.

In conclusion, this multi-omics integrative analysis exhibits advantages in the interpretation of previous GWAS results, identification of functional SNP and genetic mechanisms as well as understanding of biological processes associated with expression of beef cattle traits which could enhance genomic prediction and disease diagnosis.

Preface

This PhD thesis is an original work by Jiyuan Li. As detailed in the following, Chapter 2 of this thesis has been published in Frontiers in Genetics, Chapter 3 has been published in BMC Genomics, and Chapter 4 has been published in Scientific Reports.

All the animals used in in Chapter 2 - Chapter 4 were cared for according to the guidelines established by the Canadian Council on Animal Care (2009) and approved by local animal user committees including University of Alberta Faculty Animal Policy and Welfare Committee (2002-46B for the Bovine Genome Project and AUP00000777 for the Kinsella Breeding Project) and the AAFC Lacombe Research Center Animal Care Committee (200703 and 200807 for the Phenomic Gap Project). The study of Chapter 5 was conducted in accordance with the Canadian Council of Animal Care (2009) guidelines and recommendations (CCAC, 2009). All experimental procedures were reviewed and approved by the University of Calgary Veterinary Sciences Animal Care Committee (AC15-0109).

Chapter 2 of this thesis has been published as Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., et al. (2020). Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle. Front. Genet. 11, 538600. doi:10.3389/fgene.2020.538600. I carried out data analyses and initiated, drafted, and revised the manuscript with help from Plastow, G. S., Akanno, E. C., and Valente, T. S. Akanno, E. C. helped to manage quality control of data, construct statistic models. Valente, T. S. helped with single-step genomic BLUP analyses. Abo-Ismail, M. helped with candidate gene mapping and functional enrichment analyses. Karisa, B. K. designed the primary study, helped with data collection, and provided information on animals, blood samples and NMR spectra in the materials section. Wang, Z. contributed to the statistics and genetics background of the study. Plastow, G. S. was the

principal investigator of the project, participated in project management and experimental design. All authors read and approved the final manuscript.

Chapter 3 of this thesis has been published as Li, J., Mukiibi, R., Wang, Y., Plastow, G. S., and Li, C. (2021). Identification of candidate genes and enriched biological functions for feed efficiency traits by integrating plasma metabolites and imputed whole genome sequence variants in beef cattle. BMC Genomics 22, 823. doi:10.1186/S12864-021-08064-5. I performed data consolidation and data analyses in this study and wrote the draft of the paper. Li, C. conceived, designed, and oversaw the study. Plastow, G. S. participated in the design of the study and contributed to acquisition of data. Wang, Y. imputed 50K genotype of animals used in this study to whole genome sequence variants. Mukiibi, R. assisted with functional enrichment analyses and interpretation of IPA results. Li, C., Plastow, G. S., Mukiibi, R., and I revised the paper. All authors read, commented, and approved the final manuscript.

Chapter 4 of this thesis has been published as Li, J., Wang, Y., Mukiibi, R., Karisa, B. K., Plastow, G.S., and Li, C. (2022). Integrative analyses of genomic and metabolomic data reveal genetic mechanisms associated with carcass merit traits in beef cattle. Scientific Reports. doi: 10.1038/s41598-022-06567-z. I performed data consolidation and data analyses in this study, prepared all figures, tables and supplementary files, and wrote the draft of the manuscript. Li, C. conceived, designed, and oversaw the study. Plastow, G.S. participated in the design of the study, and Plastow, G.S. and Karisa, B. K. contributed to acquisition of data. Wang, Y. imputed 50K genotype of animals used in this study to whole genome sequence variants. Mukiibi, R. assisted with functional enrichment analyses and interpretation of IPA results. Li, C., Plastow, G.S., Wang, Y., Mukiibi, R., Karisa, B. K., and I revised the manuscript. All authors read, commented, and approved the final manuscript.

Dedication

I would like to dedicate this to my beloved parents, Mr. Li Jie and Mrs. Li Chongfang, for their love, sacrifice, and support.

To my wife, Mrs. Li Zhen, for her love, company, and encouragement.

I am truly thankful for having you all in my life.

Acknowledgements

First and foremost, I would like to express my deepest gratitude and sincere thanks to my supervisor, Prof. Graham Plastow, for his professional guidance and continuous support during the PhD program. His immense knowledge, plentiful experience, and insightful and creative ideas have inspired me in all time of my academic research and daily life.

I would especially like to thank Dr. Everestus Akanno and Dr. Zhiquan Wang as the supervisory committee members for their invaluable and constructive comments, suggestions, and criticisms. Thanks to them, I developed the scientific thinking and skills necessary to conduct research in a professional manner.

Lots of thanks to Dr. Changxi Li for his teaching and guidance when I did my studies about feed efficiency and carcass merit traits. I am also very grateful for the help of the people who contributed to my publications and thesis work. Thanks to Dr. Brian Karisa, Dr. Robert Mukiibi, Dr. Tiago Valente, and Dr. Yining Wang. Thanks to Janelle Jiminez for the help with the lab works and BRD data analysis. Additionally, this thesis would not have been possible without funding from the following funding agencies: Alberta Livestock and Meat Agency (ALMA), Alberta Innovates - Bio Solutions (AIBio), Genome Alberta and AAF project (GAB-A3GP37), the European Union's Horizon 2020 research and innovation programme under grant agreement No. 815668, the Applied Livestock Genomics Program 2014 of the Genome Alberta (project code: A222), and the AMR – One Health Consortium, funded by the Major Innovation Fund Program of the Alberta Ministry of Economic Development, Trade and Tourism.

I would like to thank my bigger family, the Livestock Gentec Center: Dr. Chunyan Zhang, Dr. Dagnachew Hailemariam, Dr. Elda Dervishi, Dr. Feng Zhang, Dr. Ghader Manafiazar, Haesu Ko, Dr. Jennifer Patterson, Dr. Kirill Krivushin, Kristin Lee, Dr. Marzieh Heidaritabar, Nahid

Bandi Kenari, Prof. Paul Stothard, Prof. Rongcai Yang, Dr. Seyed Ali Goldansaz, Dr. Shuai Zhi, Theresa Lantz, Dr. Tianfu Yang, Dr. Xuechun Bai, and Ziqi Yang. Additionally, I have met many lifelong friends in Edmonton, and you have made my study life enjoyable.

I would like to have my sincere thanks to Dr. Zhiquan Wang, who gave me the opportunity to put my first step in this journey, for his continuous support, encouragement, invaluable and constructive suggestions during all these years. Special thanks are also extended to Mrs. Yan Meng for her words and life advice.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADG: Average daily gain

AFAT: Average backfat thickness

ATAC-Seq: Assay for transposase-accessible chromatin sequencing

AUC: Area under the receiver-operator characteristic curve

BLUP: Best linear unbiased prediction

BRD: Bovine respiratory disease

BRSV: Bovine respiratory syncytial virus

BVDV: Bovine viral diarrhea virus

BW: Body weight

ChIP-Seq: Chromatin immunoprecipitation sequencing

CI: Confidence interval

CMAR: Carcass marbling score

CPM: Count per million

DE genes: Differentially expressed genes

DMI: Dry matter intake

DNAm: DNA methylation

eQTL: Gene expression quantitative trait locus

FAANG: Functional Annotation of Animal Genomes

FC: Fold change

FDR: False discovery rate

GBLUP: Genomic best linear unbiased prediction

GFBLUP: Genomic feature best linear unbiased prediction

GLM: Generalized linear model

G matrix: Genomic relationship matrix

GVCF: Genomic variant call format

GWAS: Genome-wide association studies

HCW: Hot carcass weight

IBR: Infectious bovine rhinotracheitis

IPA: Ingenuity Pathway Analysis

LD: Linkage disequilibrium

LMY: Lean meat yield

MAF: Minor allele frequency

Methyl-Seq: Methylation sequencing

mGWAS: Metabolome-genome wide association studies

mlma: Mixed linear model association

mQTL: Metabolite quantitative trait locus

MS: Mass spectrometry

MWT: Metabolic body weight

NK cell: Natural killer cell

NMR: Nuclear magnetic resonance

PCA: Principal component analysis

PI-3: Parainfluenza type 3

PLS-DA: Partial least squares discriminant analysis

Q-Q plot: Quantile-quantile plot

QTL: Quantitative trait locus

QTN: Quantitative trait nucleotide

REA: Rib eye area

REML: Restricted maximum likelihood

RFI: Residual feed intake

RLE: Relative log expression

RNA-Seq: RNA sequencing

ROC analysis: Receiver-operator characteristic analysis

SNP: Single nucleotide polymorphism

ssGBLUP: Single-step genomic best linear unbiased prediction

TCA cycle: Tricarboxylic acid cycle

TMM: Trimmed mean M values

TSS: Transcription starting site

TTS: Transcription termination site

VIP score: Variable importance in projection score

WssGBLUP: Weighted single-step genomic best linear unbiased prediction

Chapter 1. Literature review

1.1 Economically relevant traits in beef industry

The beef cattle industry is an important component of Canadian agriculture and economy. In 2020, cash receipts from all agricultural products were $72.2 billion, of which $26.3 billion originated from livestock and livestock products, including $9.1 billion from cattle and calves (Statistics Canada, 2020). Beef cattle production also supports many other related industries, such as animal health products, feed, equipment, and marketing, which add additional billions of dollars to the Canadian economy. In addition to economic profits, cattle also bring high-quality protein sources to people in Canada and around the world. About 1 million tonnes of beef (carcass weight equivalent) are produced to meet domestic meat consumption (Statistics Canada, 2020). Approximately 0.5 million tonnes of beef (carcass weight equivalent) are exported to other countries including U.S., Japan and China, which makes Canada one of the top ten beef exporters in the world (USDA). The competitiveness and sustainability of the beef industry are largely associated with feed efficiency, carcass merit and (losses due to) infectious diseases of beef cattle production.

Feed efficiency. Feed efficiency is defined as the ability of the animal to convert consumed feed nutrients into saleable beef products (Carstens and Tedeschi, 2006). Efficient animals consume less feed for the same amount of meat production (Hegarty et al., 2007). Feed efficiency is of great interest to beef production because feeding-related costs are the single largest variable expense (55% - 75%) in animal production (Ramsey et al., 2005; Ahola and Hill, 2012). A 2004 study reported that reducing feed inputs per unit of production could significantly improve profitability by 9 to 33% in beef cattle (Archer et al., 2004). Individual feed efficiency cannot be

easily measured, and thus several indices that represent feed efficiency have been developed, such as partial efficiency of gain (Kellner and Goodwin, 1913), feed conversion ratio (Brody, 1945), residual feed intake (Koch et al., 1963), and residual intake and body weight gain (Berry and Crowley, 2012).

Of the feed efficiency measures, residual feed intake (RFI) has gained popularity because it is phenotypically independent of growth and body size (Archer et al., 1999). RFI is defined as the difference between an animal's actual feed intake and expected feed intake based on its body size and growth (Koch et al., 1963). It describes the variation in feed intake that remains after the requirements for maintenance and growth have been met (Koch et al., 1963). For beef cattle, RFI values are calculated based on dry matter intake, average daily gain, and metabolic body weight over a certain test period (Nkrumah et al., 2006), which take both maintenance and growth requirements into consideration via a linear regression (Basarab et al., 2011). Efficient animals eat less than expected and have a negative or low RFI, while inefficient animals eat more than expected and have a positive or high RFI.

Selection of animals with low RFI could reduce feed costs without sacrificing the growth performance of beef cattle, and also could reduce the negative environmental impact of beef production. For example, a study has reported that steers and heifers consumed 11% less feed after two generations of selection for lower RFI, and they had similar weights and performance to their randomly mated contemporary groups (Arthur et al., 2001a). If divergent RFI lines were selected, daily feed consumption decreased by an average of 0.249 kg/day for each year of selection (Arthur et al., 2001a). Additionally, as beef cattle contribute up to 41% of the total livestock greenhouse gas emissions, this negative environmental impact may cause public concerns on beef production (Gerber et al., 2013; Morgavi et al., 2013). Selecting animals with low RFI will simultaneously

select animals with low methane emissions because of the positive genetic correlation between RFI and methane emission (Nkrumah et al., 2006; Hegarty et al., 2007). Therefore, improving feed efficiency based on RFI could decrease both production inputs and environmental footprint, which contribute to the profitability and sustainability of beef production.

Carcass merit. In addition to feed efficiency, carcass merit traits that directly influence the meat yield and quality grade also play an important role in the profitability of beef production. Yield grade (also known as cutability) is an estimate of the percent retail yield of the four primal cuts of beef including chuck, rib, loin and round, which is determined by hot carcass weight, backfat thickness, rib eye area, and percent of kidney, heart and pelvic fat. The yield grade is important to producers because it can affect animal value and the overall economic returns from the animal (Holland and Loveday, 2013). Quality grade is a composite of factors that affect palatability of meat, including the degree of marbling and degree of maturity. Improved carcass quality can enhance consumer acceptance of beef by increasing consumer satisfaction (Killinger et al., 2004). For example, sufficient marbling is important for beef tenderness, juiciness and flavor, which makes the degree of marbling in beef a primary factor determining quality grade in market age animals. Therefore, improving carcass merit traits related to meat yield and quality grade can increase the market price of beef products and total revenue.

Bovine respiratory disease. Controlling infectious diseases to keep the health of the cattle population is important for successful beef production. Bovine respiratory disease (BRD) as the most common and costly infectious disease for beef cattle in North America can cause a large economic loss in beef production. An American feedlot report showed that the direct cost of treatment of respiratory disease in feedlot cattle is USD $23.60 per case, and the total cost for treating all infected cattle (~2.29 million cattle) is estimated to be USD $51.12 million per annum

(USDA, 2011). Apart from treatment costs, cattle infected with BRD pathogens can have reduced performance, lower yield grades, increased mortality rates and may have reproductive loss (Chi et al., 2002; Snowder et al., 2006; Montgomery et al., 2009; Garcia et al., 2010), which leads to economic losses in beef production due to morbidity, mortality, prevention costs, loss of production and reduced carcass value (Griffin, 1997; Smith, 2000; Irsik et al., 2006; Montgomery et al., 2009). It was estimated that the annual economic losses from death, reduced feed efficiency, and treatment cost of BRD have ranged from USD $800 to $900 million in the U.S. beef industry (Chirase and Greene, 2001). Additionally, BRD can occur in every phase of beef production, from cow-calf to finishing. Particularly, BRD occurs most often within 4 weeks after weaning, because the weaning process is a stressful time for calves. Other stress factors ("stressors") that are commonly associated with BRD include transportation, commingling cattle from different sources, overcrowding, sudden and extreme weather change, dust, humidity, dehydration, hunger, and acute metabolic disturbances. The combination of these stressors and viral or parasitic infections can suppress the host immune system, allowing bacterial pathogens to rapidly reproduce in the upper respiratory tract (Griffin et al., 2010).

In North America, metaphylaxis and vaccination programs aimed mainly at bacterial pathogens are two major approaches to prevent and control BRD in large commercial feedlots. Metaphylaxis is defined as the mass treatment of an entire group or population of cattle with an antimicrobial to prevent and minimize an expected outbreak of disease (Ives and Richeson, 2015). Metaphylactic treatment can eliminate the already existing bacterial infections and protect those immunosuppressed and vulnerable animals against colonization and proliferation of pathogens, thereby reducing BRD-associated morbidity and mortality. Although metaphylaxis can effectively reduce BRD incidence in feedlots (Ives and Richeson, 2015), the use of mass medication has

resulted in drawbacks that include high medical costs and concerns about antimicrobial resistance associated with this procedure. Therefore, it has been suggested that mass medication should only be used on high-risk cattle where at least a BRD incidence rate of 30% would be expected (Currin and Whittier, 2000). The vaccination program is another common practice for the prevention of BRD and there are a number of commercial vaccines available against the bacterial agents and viruses associated with BRD (Bowland and Shewen, 2000; Larson and Step, 2012). However, the efficacy of these vaccines is inconsistent according to the literature (Larson and Step, 2012; Theurer et al., 2015; Chamorro and Palomares, 2020) possibly due to the polymicrobial nature of the disease. Additional research to determine the true effects of vaccination, type of vaccines, and routes of administration is needed (Chamorro and Palomares, 2020). Additionally, the viral infection is generally considered as antecedent to, or concurrent with, bacterial infection (Jericho and Langford, 1978). For example, most cattle infected by bovine respiratory syncytial virus (BRSV) are asymptomatic, but the primary infection induces loss of cilia or necrosis of bronchial and bronchiolar epithelial cells (Griffin et al., 2010). The reduced mucociliary clearance will result in buildup of fluid and cellular debris in the airways and alveoli, which provides an ideal environment for bacterial colonization. In some cases viral agents may produce a clinical syndrome consistent with BRD even though there is no bacterial co-infection (Decaro et al., 2008). Other common viral agents that are implicated in BRD, such as bovine viral diarrhea virus (BVDV), infectious bovine rhinotracheitis (IBR) and parainfluenza type 3 (PI-3) could also be observed in BRD cases. *Mannheimia haemolytica* and *Pasteurella multocida* are the most common bacterial pathogens isolated from BRD infected beef cattle (Dabo et al., 2008; Rice et al., 2008; Griffin et al., 2010; Klima et al., 2014). The co-infection of multiple viral and bacterial pathogens and non-specific symptoms of BRD make the control and diagnosis of BRD a challenge.

1.2 Genetic improvement of complex traits

Given the importance of feed efficiency, carcass merit, and resistance to BRD in determining the profitability and sustainability of beef production, researchers have been attempting to improve these important traits by genetic methods. However, the measurement of these traits is an expensive and time-consuming process (Pryce et al., 2014). For example, some carcass traits are expressed at later stages of animal production and are mostly assessed at slaughter, although some carcass traits (e.g., backfat) can be measured on live animals using real-time ultrasound imaging technologies (Schröder and Staufenbiel, 2006). Additionally, it is difficult to record and establish accurate pedigree information in commercial beef production systems in Canada unlike for dairy cows and pigs due to a low level of artificial insemination use and the high rate of crossbreeding. These factors restrict the genetic improvement of these traits using traditional genetic selection methods, such as best linear unbiased prediction (BLUP) which requires accurate phenotypic records of the individual and its relatives. With the development of genotyping technology, genetic markers associated with target breeding traits are used to assist selection (Lande and Thompson, 1990). However, target traits are usually complex and controlled by many genes with minor effects, meaning that a few genetic markers explain limited genetic variance and individually they contribute little to the genetic gain (Bernardo, 2008). In 2001, Meuwissen et al. (2001) proposed the concept of genomic selection, which is a method for predicting the genetic merit of selection candidates without phenotypes by estimating effects of all genetic markers across the whole genome. As the availability of affordable high-density genotyping services increases, genomic selection has been widely used in animal breeding and remarkably improves the selection accuracy as well as accelerates breeding progress (Hayes et al., 2009; García-Ruiz et al., 2016; Meuwissen et al., 2016; Doublet et al., 2019). Thus, genomic

selection is a desirable method of genomic evaluation and selection for complex traits. A better understanding of the genetic architecture underlying complex traits (e.g., heritability of traits, and DNA variants responsible for phenotypic variation) will help to develop a more effective genomic prediction strategy to further enhance feasibility of genomic selection in beef cattle (Hayes et al., 2009; Snelling et al., 2012).

1.3 Heritability and genome-wide association studies of traits of interest

The narrow-sense heritability of a trait refers to the proportion of phenotypic variation that is explained by additive genetic variation among individuals in a population. Heritability estimates range in value from 0 to 1. If heritability is equal to 1, then all variation in a population is due to differences or variation between genotypes. If heritability is equal to 0, all phenotypic variation in the population comes from differences in the environments experienced by individuals. The above-mentioned important traits (RFI, carcass merit traits, and susceptibility or resistance to BRD) have shown considerable variation among animals with low to moderate heritability estimates, indicating that reasonable responses could be achieved through genetic improvement. For example, the heritability of RFI has been estimated to range from 0.22 ± 0.02 to 0.68 ± 0.14 (Arthur et al., 2001; Mao et al., 2013; Ceacero et al., 2016; Torres-Vázquez et al., 2018; Zhang et al., 2020). The heritability of carcass merit traits, such as carcass weight (0.31 ± 0.04 to 0.49 ± 0.03), backfat thickness (0.26 ± 0.03 to 0.49 ± 0.05), rib eye area (0.32 ± 0.04 to 0.43 ± 0.05), and marbling score (0.33 ± 0.05 to 0.61 ± 0.06) were also estimated in different populations (Mehrban et al., 2017; Grigoletto et al., 2020; Wang et al., 2020). In addition, genetics plays a significant role in regulating immune response and determining susceptibility and resistance to BRD (Muggli-Cockett et al., 1992; Snowder et al., 2006; Snowder, 2009; Emam et al., 2019). Cattle of similar physiological characteristics and housed under the same environment show distinct individual

variability in their tendency to develop the disease and the severity of the resultant clinical signs, which implies a certain degree of genetic effect on immune response and disease susceptibility and resistance. In feedlot beef cattle, estimates of heritability for resistance to BRD ranged from 0.04 ± 0.01 to 0.08 ± 0.01 (Snowder et al., 2006). In pre-weaned calves, the estimates of heritability for resistance to BRD ranged from 0.07 ± 0.01 to 0.19 ± 0.01 (Muggli-Cockett et al., 1992; Snowder et al., 2005). Overall, these studies have shown that the variation of these traits is affected by genetic variation, which provides the possibility to improve these traits through genomic selection. Identification of DNA variants responsible for these traits will help design a better genomic prediction strategy to improve genomic selection accuracy (Hayes et al., 2009; Snelling et al., 2012).

Genome-wide association studies (GWAS) is an approach used in genetics research to associate specific genetic variations with phenotypes. The GWAS aim to understand the variation in complex traits and diseases by relating genotypes of large numbers of markers, such as single nucleotide polymorphisms (SNPs), to observed phenotypes. Over the past decade, GWAS have revolutionized research of the genetics of complex traits, as evidenced by numerous compelling associations for complex traits in humans (Hyde et al., 2016; Duncan et al., 2017; Zhao et al., 2017) and animals (Do et al., 2018; Zhang et al., 2020; Wu et al., 2021). For the traits of interest, many studies have endeavored to identify SNPs, quantitative trait loci (QTLs) or genes associated with RFI (Abo-Ismail et al., 2014; Seabury et al., 2017; Higgins et al., 2018; Zhang et al., 2020), carcass merit traits (Mehrban et al., 2017; Chang et al., 2018; Hay and Roberts, 2018; Srikanth et al., 2020; Wang et al., 2020), and BRD susceptibility/resistance (Neibergs et al., 2014; Hoff et al., 2019) by GWAS. Zhang et al. (2020) performed GWAS for RFI and identified 16 SNPs and 596 genes significantly associated with RFI. They also found that the distribution of DNA variant allele

substitution effects approximated a bell-shaped distribution while the distribution of additive genetic variances explained by single DNA variants followed a scaled inverse chi-squared distribution. These results indicate that RFI is controlled by many DNA variants with relatively small effects (Zhang et al., 2020). Similarly, Hay and Roberts (2018) implemented GWAS for fat thickness, marbling and rib eye area of beef cattle, and their results showed different SNP marker windows associated with carcass traits explained a small percentage of the genetic variance, which suggests the polygenic genetic nature of carcass traits in beef cattle. For BRD, several important genomic regions associated with susceptibility have been identified in pre-weaned Holstein calves using GWAS, and SNPs within these genomic regions were suggested to be further characterized and used for genomic selection (Neibergs et al., 2014). Especially for genes that are related to immune response, SNPs within these genes are considered to have potential effects on BRD (Casas et al., 2011). These GWAS for complex traits provided useful insights into the genetic architecture of complex traits and disease in the form of potential SNPs, structural variants and candidate genes. However, despite the clear successes of GWAS, these studies also have several limitations, including spurious associations (McClellan and King, 2010), the inability to pinpoint causal variants and genes (Boyle et al., 2017), and the lack of portability between populations (Ioannidis, 2007). In addition, GWAS usually do not provide evidence on how the variants affect downstream pathways and lead to the phenotypic variation or diseases, which troubles scientists in interpreting GWAS results and hinders the translation of GWAS findings into productive applications. In order to understand the "black box" between genotype to phenotype, researchers have turned to study the components between genotype and phenotype (Figure 1.1).

1.4 Intermediate phenotypes and omics data

Intermediate phenotype, also known as internal phenotype or endophenotype, is a quantitative biological trait that is reasonably heritable and considered to be involved in the development of an endpoint of interest (i.e., external phenotype). As shown in Figure 1.1, if a gene is considered as the start point and the associated external phenotype is considered as the end point, then a parameter lying between genome and phenome, such as transcriptome, proteome and metabolome, could be viewed as an intermediate phenotype (Preston and Weinberger, 2005; Kronenberg, 2012; Fontanesi, 2016). Since intermediate phenotypes are considered to be involved in the biological processes or pathways to external phenotypes, intermediate phenotypes may be the most valuable data for understanding the biological mechanisms of complex traits (Flint et al., 2014).

The data collected from different "omes" are collectively referred to as "multi-omics" data. Genomics, the first omics discipline to appear, focuses on the study of the entire genome, rather than the "genetics" that interrogated individual variants or single genes. As we discussed in section 1.3, GWAS is a successful approach that has been used to identify thousands of genetic variants associated with complex traits. Technologies associated with genomic data include genotype arrays, next generation sequencing for whole-genome sequencing, and exome sequencing. According to the central dogma of biology, RNA is a molecular intermediate between DNA and proteins, which are considered the primary functional read-out of DNA. Transcriptomics can qualitatively and quantitatively examine RNA levels genome-wide using probe-based arrays and RNA sequencing (RNA-Seq). After the transcriptomic level, the next level is protein. Mass spectrometry (MS) based methods have revolutionized the analysis and quantification of proteomic data. Recently, such MS methods have been adapted for high-throughput analyses of

thousands of proteins in cells or body fluids (Selevsek et al., 2015). Metabolomics is the comprehensive analysis of metabolites in a biological specimen (Clish, 2015). Metabolites are substances involved in metabolism, their levels and relative ratios reflect metabolic function and could be indicative of disease. Quantitative measures of metabolite levels in plasma and other tissues enables the discovery of novel genetic loci that regulates small molecules or their relative ratios (Kettunen et al., 2012; Shin et al., 2014). Nuclear magnetic resonance and MS are the two most commonly used techniques for generating metabolomic data. Overall, these omics data offer the possibility to study the transmission of genetic information, thereby reducing the gap between genotype and phenotype. Studies on different "omes" beyond the genome including the transcriptome for gene expression, the proteome for protein production, and the metabolome for products of metabolic functions have been conducted (Karisa et al., 2014; Tizioto et al., 2015a, 2015b; Baldassini et al., 2018; Mukiibi et al., 2018a; Fonseca et al., 2019; Blakebrough-Hall et al., 2020; Sun et al., 2020).

In respect of RFI, many studies on different intermediate phenotypes have been reported. At the transcriptome level, Mukiibi et al. (2018) performed whole transcriptome analyses for transcripts in liver because of its central physiological and metabolic functions in the body. The results suggest that reduced hepatic lipid synthesis and accumulation processes in feed efficient beef cattle may lead to a more efficient energy utilization thereby improving feed efficiency. At the metabolome level, Karisa et al. (2014) collected blood samples from beef cattle in three different periods of feeding (weeks 2, 6 and 10 in the feedlot) and identified blood metabolites (e.g., lysine, betaine, and choline etc.) were significantly associated with RFI. They also found that the metabolites associated with RFI varied over time (Karisa et al., 2014). Metabolic networks for RFI in each period showed that the cellular and molecular processes associated with RFI were

involved in energy and protein metabolism as well as the metabolism of urea and methane (Karisa et al., 2014). This study revealed the association between metabolites and RFI and potential metabolic networks, however, it does not study genetic effects on metabolite variation, which is essential information that could help identify functional SNPs associated with RFI. More recently studies at the proteome level have investigated the molecular controls of RFI in liver tissue (Baldassini et al., 2018; Fonseca et al., 2019). The differentially abundant proteins identified are mainly involved in energy metabolism, xenobiotic metabolism, vitamin metabolism, amino acid metabolism, mitochondrial function, oxygen transport, blood flow, ion transport, cell survival, microbial metabolism, biosynthesis of fatty acids, and antigen processing and presentation (Baldassini et al., 2018; Fonseca et al., 2019). These studies have broadened the knowledge of the biological mechanisms related to feed efficiency.

Regarding carcass merit traits, researchers have investigated biological mechanisms related to carcass merit traits using RNA-Seq data (Lee et al., 2014; Wang et al., 2017; Mukiibi et al., 2018b). These studies suggest that some cellular and molecular processes, including molecular transportation, lipid and fatty acid metabolism, carbohydrate metabolism, amino acid and protein metabolism, and genes involved in these functions may contribute to carcass merit traits (Lee et al., 2014; Wang et al., 2017; Mukiibi et al., 2018b). For example, Mukiibi et al. (2018b) conducted transcriptome analysis of liver tissues between 6 high and 6 low marbling Charolais steers, and identified 40 differentially expressed genes and 25 key biological processes associated with marbling. By performing the liver transcriptome analysis, they suggest that steers with more marbling were predicted to have downregulated liver lipid biosynthesis and relatively upregulated lipid concentration. The identification of these biological processes and their corresponding genes could improve our understanding of biological mechanisms associated with marbling and help

prioritize candidate genes for identification of causal gene polymorphisms responsible for the phenotypic variation. Furthermore, variation of carcass merit traits has also been reported to be associated with metabolite concentration (Connolly et al., 2019, 2020; Jeong et al., 2020). Connolly et al. (2019) reported the correlation between relative concentrations of 35 blood metabolites and carcass traits (marbling, rump fat thickness and carcass weight). Marbling was positively associated with the relative concentrations of seven metabolites (3-hydroxybutyrate, propionate, acetate, creatine, histidine, valine, and isoleucine), rump fat thickness was positively associated with glucose, leucine and lipids and negatively associated with anserine and arabinose, and carcass weight was negatively associated with 3-hydroxybutyrate. This study suggests that carcass merit traits are associated with metabolites, and the metabolome is an important intermediate phenotype that should be considered when studying the variation of carcass merit traits.

As for BRD, several transcriptomic studies were performed to investigate the gene expression variation and host response to BRD infection (Tizioto et al., 2015b; Scott et al., 2020; Sun et al., 2020; Jiminez et al., 2021). More than a thousand genes were differentially expressed in blood when animals were infected by BRD in feedlots (Scott et al., 2020; Jiminez et al., 2021). Considering multiple viral and bacterial pathogens involved in BRD infection, single-pathogen challenges in steers for three viral pathogens (BRSV, IBR, and BVDV) and three bacterial pathogens (*Mannheimia haemolytica*, *Pasteurella multocida*, and *Mycobacterium bovis*) were performed to identity common and unique differentially expressed genes to different challenge agents through RNA-Seq and transcriptomic analyses (Tizioto et al., 2015b). The results showed that many genes were expressed differentially to respond to specific pathogen challenges. They also observed some differentially expressed genes and pathways were common to all pathogen

challenges, which were primarily related to the innate immune response. This may be due to innate immune response being the first line of defense against invading pathogens and is not specific to a particular pathogen in the way that the adaptive immune response is. However, with the development of BRD, host response and related gene expression regulation may be different. A longitudinal blood transcriptomic analysis for feedlot beef cattle in Entry, Pulled and Close-out stage gives us a longitudinal view for the same individual from different time points (Sun et al., 2020). This study revealed the gene expression variations in disease development and showed that animals at Entry stage may have activated the initial response to BRD and those at Pulled stage coordinated a higher level of innate and adaptive immune responses (Sun et al., 2020). These studies have shown that host animals may regulate the immune response to defense against BRD pathogens or to respond to the damage caused by BRD pathogens through influencing the expression of certain genes.

Overall, these studies indicate that complex traits are commonly regulated through many genes and biological processes, and intermediate phenotypes play important roles in molecular processes and pathways related to external phenotypes. This knowledge not only improves the understanding of the molecular architecture of complex traits but also provides potential molecular biomarkers that could be used to identify animals with diseases or select animals with superior performance. For example, early diagnosis and appropriate treatment of BRD infected animals would enhance faster recovery and potentially reduce the negative impact of the disease on animal performance and productivity. However, most clinical signs of BRD are subjective, difficult to standardize, and nonspecific for BRD, such as fever, difficulty breathing, nasal discharge, depression, diminished or no appetite, lethargy, and coughing. Observation of clinical signs is the predominant tool for diagnosis of BRD and because of this the diagnosis of BRD can be

troublesome. By studying the intermediate phenotype, molecular biomarkers for BRD diagnosis and prediction have been proposed. Blakebrough-Hall et al. (2020) suggested that blood metabolome has a high accuracy in classifying BRD and non-BRD animals in feedlots, indicating the potential of using blood metabolome as a BRD diagnosis tool. This provides a new method to differentiate sick and healthy animals. Additionally, Sun et al. (2020b) proposed that the expression of *IFI6*, *IFIT3*, *ISG15*, *MX1*, and *OAS2* may be related to the stress level of beef cattle, and could be used as biomarkers to predict and recognize sick cattle when animals enter the feedlot. Although these molecular biomarkers are promising and need further validation, they provide alternative methods of identifying or predicting sick animals in feedlot production.

Despite these achievements, most of the studies only focused on the relationship between a single omics layer and the phenotype and overlooked the interrelationship between different omics layers. For example, some metabolites have been reported to be heritable and their variations are affected by genetic variants (Buitenhuis et al., 2013; Li et al., 2020), and changes in gene expression are also affected by genetic regulation (e.g., expression QTL; eQTL) (Cookson et al., 2009). However, studies of a single omics layer are difficult to identify the interrelationship between different omics layers and causal SNPs or genes. Additionally, most of the biological processes involve more than one type of biomolecule, and hence operate not solely at the level of either genome, transcriptome, proteome, or metabolome (Haas et al., 2017). Although many important biological processes associated with traits of interest have been identified, a single omics layer can only carry part of the biological information and such single omics studies could not generate a whole picture of biological and molecular background of complex traits. Therefore, with single omics research it is still difficult to answer the two important questions of GWAS as discussed in section 1.2: “Which genetic variants drive the phenotypic variation?” and “How do

genetic variants lead to phenotypic variation?". To answer these questions, it is imperative to take an integrative approach that combines multi-omics data.

1.5 Multi-omics analysis to identify functional SNPs and genes and biological functions

With the advent of high throughput sequencing and mass spectrometry technologies, it has been possible to collect two or more omics datasets at the same time and study them under a given circumstance. This multi-omics analysis could provide new insights into the interrelationships between different omics layers and help to identify functional or causal SNPs and biological functions related to traits of interest. The advantage of such multi-omics approach is that they can better answer the two above-mentioned questions of GWAS.

"Which genetic variants drive the phenotypic variation?" The resolution of GWAS is relatively low since it only requires correlation to phenotypes by neighbouring genetic markers in linkage disequilibrium (LD). For the great majority of identified genetic markers or genes, more effort is still needed to refine their identities. Multi-omics analysis has been used to pinpoint the functional or causal SNPs and genes that are associated with complex traits in both humans (Shin et al., 2014; Frost and Amos, 2018; Sun et al., 2018; Schlosser et al., 2020) and some livestock species (Fang et al., 2017b; Li et al., 2019; Xu et al., 2019; Fu et al., 2020). For example, Schlosser et al. (2020) analyzed the urinary concentrations of 1172 metabolites in 1627 individuals with reduced kidney function and they identified 240 genomic regions associated with urinary concentration of metabolites. The colocalization analysis of identified metabolite QTLs (mQTLs) and genetic association signals for human diseases found numerous significant overlaps, such as overlapped association signals at the *ALPL* locus between urinary phosphoethanolamine, urolithiasis, and kidney stones (Schlosser et al., 2020). In another human study, a comprehensive genetic association study of metabolites in blood identified 145 mQTLs, and 41 (28.3%) of these

mQTLs were reported to be associated with other human complex traits and diseases (Shin et al., 2014). These studies suggest that these overlapping QTLs may be more important than others because they may contain biologically relevant genes associated with phenotypes. In beef cattle, a pilot study combining host genotype and rumen microbiota was performed to investigate the association between SNPs, rumen microbiota and feed efficiency. The results showed that some rumen microbial features are heritable, and the heritable elements of these microbial features are affected by host additive genetics (Li et al., 2019). They also found that those heritable rumen microbial features were associated with host feed efficiency and differences in rumen volatile fatty acids. Some SNPs were identified as associated with both rumen microbiota and feed efficiency (Li et al., 2019), indicating that these overlapping SNPs may influence feed efficiency by influencing rumen microbiota. Compared with GWAS, which is based solely on the association between genotype and phenotype, this multi-omics analysis could help to refine or prioritize the real functional or causal SNPs and genes.

"How genetic variants lead to phenotypic variation?" Multi-omics analysis can assess the information flow from one omics level to the other and the interaction between different omics levels, which could provide new insights and perspectives into how genetic variants or genes influence phenotypes, and the molecular mechanisms underlying traits of interest. The biological information obtained by analyzing multi-omics data could further help to elucidate potential causal variants and genes that lead to the variation of phenotype (Hasin et al., 2017; Zhao et al., 2020). For example, Widmann et al. (2013) combined genetic, metabolomic and physiological data to dissect genes and molecular pathways that modulate differential growth at the onset of puberty. They found GnRH (Gonadotropin-releasing hormone) signaling was associated with divergent growth at the onset of puberty, and revealed two highly connected genes, *BTC* and *DGKH*, within

the GnRH network. They also indicate the functional role of an interaction network for *NCAPG* in divergent growth (Widmann et al., 2013). In another study using multi-omics data to study puberty in cattle, Cánovas et al. (2014) characterized the transcriptome of five reproductive tissues (i.e., hypothalamus, pituitary gland, ovary, uterus, and endometrium) and three puberty related tissues (i.e., longissimus dorsi muscle, adipose, and liver) of pre- and post-pubertal Brangus heifers with genotypes to study gene and molecular regulatory networks for puberty. In this study, 1,515 differentially expressed, and 943 tissue-specific genes and key transcriptional regulators were identified. Twenty-five loci containing a SNP associated with fertility traits were determined through combining the results of GWAS and RNA-Seq (Cánovas et al., 2014). Finally, they constructed pre- and post-puberty co-expression gene networks (an undirected graph to indicate which genes have a tendency to show a coordinated expression pattern across a group of samples) by combining the results from GWAS, RNA-Seq, and bovine transcription factors to reveal the changes in biomolecules and biological functions of animals before and after puberty. Overall, the multi-omics analysis has shown great potential in revealing potential regulatory mechanisms that cannot be captured by single omics data, as well as discovering genes that contain biologically relevant SNPs, which could be used in genomic selection and may lead to improved performance of genomic selection in beef cattle.

1.6 Functional information obtained through multi-omics analysis could enhance genomic selection

Genomic selection is a method of choice for selective breeding and improvement of economically relevant traits in livestock. One major factor that affects the rate of genetic improvement is the accuracy of genomic prediction of selection candidates (Georges et al., 2018). Genomic prediction accuracy, the correlation between predicted and true breeding values, depends

on the LD between DNA marker and QTLs, the size of the training population, the genetic relationship among animals in the training and validation populations, number of DNA markers in the evaluation panel, genetic architecture of the trait, and the statistical method used for the genetic evaluation (Habier et al., 2007; Goddard, 2009; Goddard and Hayes, 2009; Zhang et al., 2019). In beef cattle, genomic prediction has been applied to feed efficiency and carcass merit traits (Akanno et al., 2014; Lourenco et al., 2015; Fernandes Júnior et al., 2016; Lu et al., 2016). Nevertheless, the accuracy of genomic prediction for these traits remains relatively low (Mujibi et al., 2011; Bolormaa et al., 2013; Akanno et al., 2014; Lu et al., 2016; Silva et al., 2016). The lower prediction accuracy of beef cattle may be due to the lack of a large number of reference populations within a breed, which is caused by the high rate of crossbreeding of beef cattle and the high cost of recording certain traits (e.g., feed efficiency). Meanwhile, information from another breed is much less useful than information from the target breed because animals of different breeds share much smaller chromosome segments than animals of the same breed (Bolormaa et al., 2013; Chen et al., 2013; Akanno et al., 2014). Additionally, the polygenic nature of these complex traits and insufficient understanding of causal DNA variants/genes and biological mechanisms are also important reasons for low accuracy (Snelling et al., 2013; Fang et al., 2017b; Wang et al., 2020; Zhang et al., 2020).

The low accuracy of genomic prediction in beef cattle makes it difficult to identify true animals with high breeding values, which affects the rate of genetic gain in beef cattle. Consequently, multiple strategies including different analysis methods, SNP densities, and reference population construction strategies have been evaluated with the objective of improving genomic prediction accuracy in beef cattle. For example, with respect to feed efficiency, Lu et al. (2016) tested the impact of genomic distance between training and validation population, training

population size, statistical methods (genomic BLUP and Bayes C), density of genetic markers (50K and imputed high density genotype) on prediction accuracy for feed efficiency traits in multibreed and crossbred beef cattle. They found that the Bayes C could obtain marginally higher accuracy than genomic BLUP, and adding animals that were less related to the training population could not increase the prediction accuracy (Lu et al., 2016). According to the comparison of prediction accuracy between 50K and imputed high density genotype, they found that 50K genotypes are more effective for predicting genomic estimate breeding value in purebred cattle while imputed high density genotypes found utility when dealing with composites and crossbreds (Lu et al., 2016). Chen et al. (2013) evaluated the impact of the reference population construction strategies for feed efficiency in two purebred (Angus and Charolais) populations. The results showed that the accuracy of the prediction for Angus and Charolais reduced dramatically (0.53 to 0.16 and 0.64 to 0.10, respectively) when the reference population used for genomic prediction changed from within breed to across breed. This study indicates that genetic relationship of selection candidates with the reference population has a greater impact on the prediction accuracy. In another study, phenotype and genotype records of cattle from multiple sources were combined as a reference population to improve the genomic prediction accuracy for feed efficiency (Khansefid et al., 2014). The results showed that the multibreed reference population increased the accuracy of genomic prediction slightly by an average of 5% (Khansefid et al., 2014). However, due to the inconsistent LD across breeds of the beef population and the small number of reference animals within a breed (De Roos et al., 2009; Meuwissen et al., 2016), the improvement of prediction accuracy is relatively small. Therefore, achieving a reasonably greater accuracy of genomic prediction in beef cattle remains a challenge.

One strategy to improve prediction accuracy is to include biological or functional information or causal DNA variants into the SNP panels or in the statistical models (Snelling et al., 2013; MacLeod et al., 2016; Fang et al., 2017b). This could provide an opportunity to improve genomic prediction through reducing reliance on the LD for gene or regulatory SNP markers to capture the effects of QTLs (Snelling et al., 2013). The functionally enriched SNP panels are expected to result in a higher genomic prediction accuracy for crossbred animals because prediction would be based on the functional SNP effects expected to be relatively stable across breeds or populations, rather than LD that is usually unstable across population or usually broken down through crossing (Snelling et al., 2013). Several pilot studies have been conducted to incorporate functional SNPs and biological information into genomic prediction models or SNP evaluation panels (Melzer et al., 2013; MacLeod et al., 2016; Sarup et al., 2016; Fang et al., 2017a; Gebreyesus et al., 2019). For example, Fang et al. (2017a) applied an extended genomic BLUP (GBLUP) model called genomic feature BLUP (GFBLUP) that includes a separate random effect for the joint action of SNPs within genomic feature which are obtained from RNA differential expression analyses. Compared to GBLUP, the accuracy of genomic prediction for mastitis and milk production traits with GFBLUP was marginally improved (3.2% to 3.9%) in within-breed prediction but significantly increased (164.4%) in across-breed prediction (Fang et al., 2017a). Theoretically, the genomic features could be defined from various sources of biological knowledge (e.g., metabolomics and proteomics) and the GFBLUP model could be applied to other complex traits. MacLeod et al. (2016) introduced Bayes RC based on Bayes R that incorporates prior biological information in the analysis. The information can be derived from a range of sources, including variant annotation, candidate genes and known causal variants, and this information is then incorporated objectively in the analysis based on the evidence of enrichment in the data. The

result showed that Bayes RC could increase the power to detect causal variants and increase the accuracy of genomic prediction. The relative improvement for genomic prediction was most apparent in validation populations that were not closely related to the training population (MacLeod et al., 2016). These studies have shown the advantage of including functional information in the genomic prediction for crossbred animals (commercial crossbred beef cattle as found in the Canadian industry) or less related populations (the validation population is not closely related to training population). Therefore, as biological knowledge accumulates regarding functional regions of the genome for a range of traits and species, approaches such as GFBLUP and Bayes RC will become increasingly useful in genomic prediction. The multi-omics analysis as a promising approach to identify more reliable and accurate functional or causal SNPs and genes would also contribute to the improvement of genomic prediction accuracy of complex traits in beef cattle.

In addition to improve genomic prediction accuracy, including functional or causal DNA variants into the SNP panels may decrease the DNA marker density used in genomic prediction while retaining accuracy. This may reduce the cost of genotyping services during the genome selection process. Melzer et al. (2013) integrated genomic and metabolomic data into genomic prediction of three traditional milk traits in dairy cows. In that study, regression methods were applied to identify important milk metabolites and then those SNPs with significant genetic effects on important metabolites were identified and used to predict milk traits. Compared with the classical approach that uses all SNPs (40,317) in prediction, the panel containing metabolite associated SNPs could achieve similar prediction precision with less than 1% of the total amount of SNPs (Melzer et al., 2013). Although some promising results have been achieved, we would like to highlight that this is not always the case. In another study, a 50K functional SNP panel in

which SNPs were located within or close to the differentially expressed genes associated with feed efficiency traits from their own studies and other published literature was assembled, and its prediction accuracy was compared with prediction accuracy of a commercial 50K SNP panel and a random panel (Mukiibi, 2019). The results showed that the genomic prediction accuracies of the three SNP panels were similar for all the traits under all the genomic prediction scenarios (within-breed or across breeds) when the standard errors were considered, indicating that the functional panel did not lead to tangible improvement in genomic prediction for the feed efficiency and the related traits in the beef cattle populations investigated. One of the reasons is that SNPs in the functional panel are not real functional SNPs, and exact causal SNPs might be still unknown. This indicates that identifying functional or causal SNPs and genes just based on the correlation between a single omics level and phenotype is insufficient and may cause either false positive or false negative results. Including unrelated SNPs in genomic prediction may just add more background noise and create more prediction errors (Li et al., 2018). This also highlights the importance and necessity of employing multiple omics tools to identify or refine causal SNPs associated with complex traits. In addition, the biological information carried by multi-omics data can assist genomic feature preselection or prioritize SNPs used in genomic prediction models or evaluation SNP panels, and can further contribute to improving performance of genomic prediction (Ye et al., 2020).

In summary, multi-omics analysis is becoming increasingly important to bridge the gap from genotype to phenotype and holistically study complex traits. However, for the traits under consideration in the current study (i.e., feed efficiency, carcass merit, and BRD), the successful application of such multi-omics analysis is still rare. It is necessary to conduct multi-omics studies on these traits to improve the understanding of their genetic and molecular background.

Additionally, in view of the relatively low accuracy of genomic prediction for feed efficiency, carcass merit traits and BRD, identification of more reliable and accurate functional or causal SNPs and genes using multi-omics analysis is also a promising way to improve genomic prediction for these traits.

1.7 Research objectives

The overall goal of this thesis was to utilise a multi-omics dataset (genomics, transcriptomics, and metabolomics) to study the genetic and molecular background of feed efficiency and carcass merit traits as well as bovine respiratory disease of beef cattle, and to identify functional or causal SNPs and genes associated with these traits. Specific objectives included:

1) Exploring the genetic architecture of blood metabolites in beef cattle
2) Performing integrative analyses of genomic and metabolomic data to reveal genetic mechanisms underlying feed efficiency traits in beef cattle
3) Performing integrative analyses of genomic and metabolomic data to reveal genetic mechanisms underlying carcass merit traits in beef cattle
4) Exploring the genetic background and molecular mechanisms of bovine respiratory disease using genomic and transcriptomic data in feedlot beef cattle

1.8 References

Abo-Ismail, M. K., Vander Voort, G., Squires, J. J., Swanson, K. C., Mandell, I. B., Liao, X., et al. (2014). Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. *BMC Genet.* 15, 14. doi:10.1186/1471-2156-15-14.

Ahola, J. K., and Hill, R. A. (2012). "Input Factors Affecting Profitability: A Changing Paradigm and a Challenging Time," in *Feed Efficiency in the Beef Industry* (Hoboken, New Jersey: Wiley-Blackwell), 7–19. doi:10.1002/9781118392331.ch1.

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits

of beef cattle - An independent validation study. *J. Anim. Sci.* 92, 2896–2904. doi:10.2527/jas.2013-7374.

Archer, J. A., Barwick, S. A., and Graser, H. U. (2004). Economic evaluation of beef cattle breeding schemes incorporating performance testing of young bulls for feed intake. *Aust. J. Exp. Agric.* 44, 393–404. doi:10.1071/ea02054.

Archer, J. A., Richardson, E. C., Herd, R. M., and Arthur, P. F. (1999). Potential for selection to improve efficiency of feed use in beef cattle: A review. *Aust. J. Agric. Res.* 50, 147–161. doi:10.1071/a98075.

Arthur, P. F., Archer, J. A., Herd, R. M., and Melville, G. J. (2001a). Response To Selection For Net Feed Intake in Beef Cattle. in *Proceedings of the Association for the Advancement of Animal Breedings and Genetics*, 135–138.

Arthur, P. F., Archer, J. A., Johnston, D. J., Herd, R. M., Richardson, E. C., and Parnell, P. F. (2001b). Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *J. Anim. Sci.* 79, 2805. doi:10.2527/2001.79112805x.

Baldassini, W. A., Bonilha, S. F. M., Branco, R. H., Vieira, J. C. S., Padilha, P. M., and Lanna, D. P. D. (2018). Proteomic investigation of liver from beef cattle (Bos indicus) divergently ranked on residual feed intake. *Mol. Biol. Rep.* 45, 2765–2773. doi:10.1007/s11033-018-4341-2.

Bernardo, R. (2008). Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Sci.* 48, 1649–1664. doi:10.2135/CROPSCI2008.03.0131.

Berry, D. P., and Crowley, J. J. (2012). Residual intake and body weight gain: A new measure of efficiency in growing cattle. *J. Anim. Sci.* 90, 109–115. doi:10.2527/JAS.2011-4245.

Blakebrough-Hall, C., Dona, A., D'occhio, M. J., McMeniman, J., and González, L. A. (2020). Diagnosis of Bovine Respiratory Disease in feedlot cattle using blood 1H NMR metabolomics. *Sci. Rep.* 10, 115. doi:10.1038/s41598-019-56809-w.

Bolormaa, S., Pryce, J. E., Kemper, K., Savin, K., Hayes, B. J., Barendse, W., et al. (2013). Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in Bos taurus, Bos indicus, and composite beef cattle. *J. Anim. Sci.* 91, 3088–3104. doi:10.2527/JAS.2012-5827.

Bowland, S. L., and Shewen, P. E. (2000). Bovine respiratory disease: Commercial vaccines currently available in Canada. *Can. Vet. J.* 41, 33–48.

Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169, 1177. doi:10.1016/J.CELL.2017.05.038.

Brody, S. (1945). *Bioenergetics and growth; with special reference to the efficiency complex in domestic animals*. New York, USA: Reinhold Publishing Corp Available at: https://psycnet.apa.org/record/1946-03094-000.

Buitenhuis, A. J. J., Sundekilde, U. K. K., Poulsen, N. A. A., Bertram, H. C. C., Larsen, L. B. B., and Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *J. Dairy Sci.* 96, 3285–3295. doi:10.3168/jds.2012-5914.

Cánovas, A., Reverter, A., DeAtley, K. L., Ashley, R. L., Colgrave, M. L., Fortes, M. R. S., et al. (2014). Multi-Tissue Omics Analyses Reveal Molecular Regulatory Networks for Puberty in Composite Beef Cattle. *PLoS One* 9, e102551. doi:10.1371/JOURNAL.PONE.0102551.

Carstens, G., and Tedeschi, L. (2006). Defining feed efficiency in beef cattle. in *Proceedings of Beef Improvement Federation 38th Annual Research Symposium and Annual Meeting*, 12–21.

Casas, E., Garcia, M. D., Wells, J. E., and Smith, T. P. L. (2011). Association of single nucleotide polymorphisms in the ANKRA2 and CD180 genes with bovine respiratory disease and presence of Mycobacterium avium subsp. paratuberculosis1. *Anim. Genet.* 42, 571–577. doi:10.1111/j.1365-2052.2011.02189.x.

Ceacero, T. M., Mercadante, M. E. Z., Cyrillo, J. N. D. S. G., Canesin, R. C., Bonilha, S. F. M., and De Albuquerque, L. G. (2016). Phenotypic and genetic correlations of feed efficiency traits with growth and carcass traits in nellore cattle selected for postweaning weight. *PLoS One* 11, e0161366. doi:10.1371/journal.pone.0161366.

Chamorro, M. F., and Palomares, R. A. (2020). Bovine Respiratory Disease Vaccination Against Viral Pathogens: Modified-Live Versus Inactivated Antigen Vaccines, Intranasal Versus Parenteral, What Is the Evidence? *Vet. Clin. North Am. Food Anim. Pract.* 36, 461. doi:10.1016/J.CVFA.2020.03.006.

Chang, T., Xia, J., Xu, L., Wang, X., Zhu, B., Zhang, L., et al. (2018). A genome-wide association study suggests several novel candidate genes for carcass traits in Chinese Simmental beef cattle. *Anim. Genet.* 49, 312–316. doi:10.1111/AGE.12667.

Chen, L., Schenkel, F., Vinsky, M., Crews, D. H., and Li, C. (2013). Accuracy of predicting genomic breeding values for residual feed intake in Angus and Charolais beef cattle. *J. Anim. Sci.* 91, 4669–4678. doi:10.2527/JAS.2013-5715.

Chi, J., VanLeeuwen, J. A., Weersink, A., and Keefe, G. P. (2002). Direct production losses and treatment costs from bovine viral diarrhoea virus, bovine leukosis virus, Mycobacterium avium subspecies paratuberculosis, and Neospora caninum. *Prev. Vet. Med.* 55, 137–153. doi:10.1016/S0167-5877(02)00094-6.

Chirase, N. K., and Greene, L. W. (2001). Dietary zinc and manganese sources administered from the fetal stage onwards affect immune response of transit stressed and virus infected offspring steer calves. *Anim. Feed Sci. Technol.* 93, 217–228. doi:10.1016/S0377-8401(01)00277-2.

Clish, C. B. (2015). Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb. Mol. Case Stud.* 1, a000588. doi:10.1101/MCS.A000588.

Connolly, S., Dona, A., Hamblin, D., D'Occhio, M. J., and González, L. A. (2020). Changes in the blood metabolome of Wagyu crossbred steers with time in the feedlot and relationships with marbling. *Sci. Rep.* 10, 18987. doi:10.1038/s41598-020-76101-6.

Connolly, S., Dona, A., Wilkinson-White, L., Hamblin, D., D'Occhio, M., and González, L. A. (2019). Relationship of the blood metabolome to subsequent carcass traits at slaughter in feedlot Wagyu crossbred steers. *Sci. Rep.* 9, 15139. doi:10.1038/s41598-019-51655-2.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex

disease traits with global gene expression. *Nat. Rev. Genet.* 10, 184–194. doi:10.1038/nrg2537.

Dabo, S. M., Taylor, J. D., and Confer, A. W. (2008). Pasteurella multocida and bovine respiratory disease. *Anim. Heal. Res. Rev.* 8, 129–150. doi:10.1017/S1466252307001399.

De Roos, A. P. W., Hayes, B. J., and Goddard, M. E. (2009). Reliability of genomic predictions across multiple populations. *Genetics* 183, 1545–1553. doi:10.1534/genetics.109.104935.

Decaro, N., Campolo, M., Desario, C., Cirone, F., D'abramo, M., Lorusso, E., et al. (2008). Respiratory Disease Associated with Bovine Coronavirus Infection in Cattle Herds in Southern Italy. *J. Vet. Diagnostic Investig.* 20, 28–32. doi:10.1177/104063870802000105.

Do, D. N., Schenkel, F. S., Miglior, F., Zhao, X., and Ibeagha-Awemu, E. M. (2018). Genome wide association study identifies novel potential candidate genes for bovine milk cholesterol content. *Sci. Rep.* 8, 13239. doi:10.1038/s41598-018-31427-0.

Doublet, A. C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., et al. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genet. Sel. Evol.* 51, 52. doi:10.1186/S12711-019-0495-1/FIGURES/7.

Duncan, L., Yilmaz, Z., Walters, R., Goldstein, J., Anttila, V., Bulik-Sullivan, B., et al. (2017). Genome-Wide Association Study Reveals First Locus for Anorexia Nervosa and Metabolic Correlations. *Am. J. Psychiatry* 174, 850. doi:10.1176/APPI.AJP.2017.16121402.

Emam, M., Livernois, A., Paibomesai, M., Atalla, H., and Mallard, B. (2019). Genetic and Epigenetic Regulation of Immune Response and Resistance to Infectious Diseases in Domestic Ruminants. *Vet. Clin. North Am. - Food Anim. Pract.* 35, 405–429. doi:10.1016/j.cvfa.2019.07.002.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017a). Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle by mapping variants to hepatic transcriptomic regions responsive to intra-mammary infection. *Genet. Sel. Evol.* 49, 44. doi:10.1186/S12711-017-0319-0.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017b). Use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits within and between dairy cattle breeds. *BMC Genomics* 18, 604. doi:10.1186/S12864-017-4004-Z.

Fernandes Júnior, G. A., Rosa, G. J. M., Valente, B. D., Carvalheiro, R., Baldi, F., Garcia, D. A., et al. (2016). Genomic prediction of breeding values for carcass traits in Nellore cattle. *Genet. Sel. Evol.* 48, 7. doi:10.1186/s12711-016-0188-y.

Flint, J., Timpson, N., and Munafò, M. (2014). Assessing the utility of intermediate phenotypes for genetic mapping of psychiatric disease. *Trends Neurosci.* 37, 733. doi:10.1016/J.TINS.2014.08.007.

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., et al. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Sci. Rep.* 9, 5364. doi:10.1038/s41598-019-41813-x.

Fontanesi, L. (2016). "Merging Metabolomics, Genetics, and Genomics in Livestock to Dissect Complex Production Traits," in *Systems Biology in Animal Production and Health, Vol. 1*

(Springer, Cham), 43–62. doi:10.1007/978-3-319-43335-6_3.

Frost, H. R., and Amos, C. I. (2018). A multi-omics approach for identifying important pathways and genes in human cancer. *BMC Bioinformatics* 19, 479. doi:10.1186/S12859-018-2476-8.

Fu, Y., Wang, L., Tang, Z., Yin, D., Xu, J., Fan, Y., et al. (2020). An integration analysis based on genomic, transcriptomic and QTX information reveals credible candidate genes for fat-related traits in pigs. *Anim. Genet.* 51, 683–693. doi:10.1111/AGE.12971.

Garcia, M. D., Thallman, R. M., Wheeler, T. L., Shackelford, S. D., and Casas, E. (2010). Effect of bovine respiratory disease and overall pathogenic disease incidence on carcass traits. *J. Anim. Sci.* 88, 491–496. doi:10.2527/jas.2009-1874.

García-Ruiz, A., Cole, J. B., VanRaden, P. M., Wiggans, G. R., Ruiz-López, F. J., and Tassell, C. P. Van (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci.* 113, E3995–E4004. doi:10.1073/PNAS.1519061113.

Gebreyesus, G., Bovenhuis, H., Lund, M. S., Poulsen, N. A., Sun, D., and Buitenhuis, B. (2019). Reliability of genomic prediction for milk fatty acid composition by using a multi-population reference and incorporating GWAS results. *Genet. Sel. Evol.* 51, 16. doi:10.1186/S12711-019-0460-Z.

Georges, M., Charlier, C., and Hayes, B. (2018). Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* 20, 135–156. doi:10.1038/s41576-018-0082-2.

Gerber, P. J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., et al. (2013). *Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities*. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).

Goddard, M. (2009). Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica* 136, 245–257. doi:10.1007/s10709-008-9308-0.

Goddard, M. E., and Hayes, B. J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet.* 10, 381–391. doi:10.1038/nrg2575.

Griffin, D. (1997). Economic impact associated with respiratory disease in beef cattle. *Vet. Clin. North Am. Food Anim. Pract.* 13, 367–377. doi:10.1016/S0749-0720(15)30302-9.

Griffin, D., Chengappa, M. M., Kuszak, J., and McVey, D. S. (2010). Bacterial pathogens of the bovine respiratory disease complex. *Vet. Clin. North Am. - Food Anim. Pract.* 26, 381–394. doi:10.1016/j.cvfa.2010.04.004.

Grigoletto, L., Ferraz, J. B. S., Oliveira, H. R., Eler, J. P., Bussiman, F. O., Abreu Silva, B. C., et al. (2020). Genetic Architecture of Carcass and Meat Quality Traits in Montana Tropical® Composite Beef Cattle. *Front. Genet.* 0, 123. doi:10.3389/FGENE.2020.00123.

Haas, R., Zelezniak, A., Iacovacci, J., Kamrad, S., Townsend, S. J., and Ralser, M. (2017). Designing and interpreting “multi-omic” experiments that may change our understanding of biology. *Curr. Opin. Syst. Biol.* 6, 37–45. doi:10.1016/j.coisb.2017.08.009.

Habier, D., Fernando, R. L., and Dekkers, J. C. M. (2007). The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics* 177, 2389–2397.

doi:10.1534/GENETICS.107.081190.

Hasin, Y., Seldin, M., and Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biol.* 18, 1–15. doi:10.1186/s13059-017-1215-1.

Hay, E. H., and Roberts, A. (2018). Genome-wide association study for carcass traits in a composite beef cattle breed. *Livest. Sci.* 213, 35–43. doi:10.1016/J.LIVSCI.2018.04.018.

Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K., and Goddard, M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41, 51. doi:10.1186/1297-9686-41-51/FIGURES/2.

Hayes, B. J., Visscher, P. M., and Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res. (Camb).* 91, 47–60. doi:10.1017/S0016672308009981.

Hegarty, R. S., Goopy, J. P., Herd, R. M., and McCorkell, B. (2007). Cattle selected for lower residual feed intake have reduced daily methane production. *J. Anim. Sci.* 85, 1479–1486. doi:10.2527/jas.2006-236.

Higgins, M. G., Fitzsimons, C., McClure, M. C., McKenna, C., Conroy, S., Kenny, D. A., et al. (2018). GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle. *Sci. Rep.* 8, 14301. doi:10.1038/s41598-018-32374-6.

Hoff, J. L., Decker, J. E., Schnabel, R. D., Seabury, C. M., Neibergs, H. L., and Taylor, J. F. (2019). QTL-mapping and genomic prediction for bovine respiratory disease in U.S. Holsteins using sequence imputation and feature selection. *BMC Genomics* 20, 555. doi:10.1186/s12864-019-5941-5.

Holland, R., and Loveday, D. (2013). *Understanding Yield Grades and Quality Grades for Value-Added Beef Producers and Marketers*. Institute of Agriculture, University of Tennessee, Knoxville.

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., et al. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* 48, 1031–1036. doi:10.1038/NG.3623.

Irsik, M., Langemeier, M., Schroeder, T., Spire, M., and Deen Roder, J. (2006). Estimating the Effects of Animal Health on the Performance of Feedlot Cattle. *Bov. Pract.* 40, 65–74. doi:org/10.21423/bovine-vol40no2p65-74.

Ives, S. E., and Richeson, J. T. (2015). Use of Antimicrobial Metaphylaxis for the Control of Bovine Respiratory Disease in High-Risk Cattle. *Vet. Clin. North Am. - Food Anim. Pract.* 31, 341–350. doi:10.1016/j.cvfa.2015.05.008.

Jeong, J. Y., Kim, M., Ji, S. Y., Baek, Y. C., Lee, S., Oh, Y. K., et al. (2020). Metabolomics analysis of the beef samples with different meat qualities and tastes. *Food Sci. Anim. Resour.* 40, 924–937. doi:10.5851/KOSFA.2020.E59.

Jericho, K. W. F., and Langford, E. V. (1978). Pneumonia in Calves Produced with Aerosols of Bovine Herpesvirus 1 and Pasteurella haemolytica. *Can. J. Comp. Med.* 42, 269.

Jiminez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-

blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front. Genet.* 12, 257. doi:10.3389/fgene.2021.627623.

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., et al. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livest. Sci.* 165, 200–211. doi:10.1016/j.livsci.2014.03.002.

Kellner, O. J., and Goodwin, W. (1913). *The Scientific Feeding of Animals*. London, UK: Duckworth & Co. Available at: https://books.google.ca/books?id=iv01AQAAIAAJ&dq=The+Scientific+Feeding+of+Animals&lr=&source=gbs_navlinks_s.

Kettunen, J., Tukiainen, T., Sarin, A. P., Ortega-Alonso, A., Tikkanen, E., Lyytikäinen, L. P., et al. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* 44, 269–276. doi:10.1038/ng.1073.

Khansefid, M., Pryce, J. E., Bolormaa, S., Miller, S. P., Wang, Z., Li, C., et al. (2014). Estimation of genomic breeding values for residual feed intake in a multibreed cattle population. *J. Anim. Sci.* 92, 3270–3283. doi:10.2527/JAS.2014-7375.

Killinger, K. M., Calkins, C. R., Umberger, W. J., Feuz, D. M., and Eskridge, K. M. (2004). Consumer sensory acceptance and value for beef steaks of similar tenderness, but differing in marbling level. *J. Anim. Sci.* 82, 3294–3301. doi:10.2527/2004.82113294X.

Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., et al. (2014). Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *J. Clin. Microbiol.* 52, 438–448. doi:10.1128/JCM.02485-13.

Koch, R. M., Swiger, L. A., Chambers, D., and Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *J. Anim. Sci.* 22, 486–494. doi:10.2527/jas1963.222486x.

Kronenberg, F. (2012). "Metabolic Traits as Intermediate Phenotypes," in *Genetics Meets Metabolomics: From Experiment to Systems Biology* (Springer, New York, NY), 255–264. doi:10.1007/978-1-4614-1689-0_15.

Lande, R., and Thompson, R. (1990). Efficiency of Marker-Assisted Selection in the Improvement of Quantitative Traits. *Genetics* 124, 756. doi:10.1093/genetics/124.3.743.

Larson, R. L., and Step, D. L. (2012). Evidence-Based Effectiveness of Vaccination Against Mannheimia haemolytica, Pasteurella multocida, and Histophilus somni in Feedlot Cattle for Mitigating the Incidence and Effect of Bovine Respiratory Disease Complex. *Vet. Clin. North Am. Food Anim. Pract.* 28, 97–106. doi:10.1016/j.cvfa.2011.12.005.

Lee, H. J., Park, H. S., Kim, W., Yoon, D., and Seo, S. (2014). Comparison of metabolic network between muscle and intramuscular adipose tissues in Hanwoo beef cattle using a systems biology approach. *Int. J. Genomics* 2014, 679437. doi:10.1155/2014/679437.

Li, B., Zhang, N., Wang, Y.-G., George, A. W., Reverter, A., and Li, Y. (2018). Genomic Prediction of Breeding Values Using a Subset of SNPs Identified by Three Machine Learning Methods. *Front. Genet.* 0, 237. doi:10.3389/FGENE.2018.00237.

Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., et al. (2019). Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in

cattle. *Microbiome* 7, 1–17. doi:10.1186/s40168-019-0699-1.

Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., et al. (2020). Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle. *Front. Genet.* 11, 538600. doi:10.3389/fgene.2020.538600.

Ioannidis, J. P. A. (2007). Non-Replication and Inconsistency in the Genome-Wide Association Setting. *Hum. Hered.* 64, 203–213. doi:10.1159/000103512.

Lourenco, D. A. L., Tsuruta, S., Fragomeni, B. O., Masuda, Y., Aguilar, I., Legarra, A., et al. (2015). Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J. Anim. Sci.* 93, 2653–2662. doi:10.2527/jas.2014-8836.

Lu, D., Akanno, E. C., Crowley, J. J., Schenkel, F., Li, H., De Pauw, M., et al. (2016). Accuracy of genomic predictions for feed efficiency traits of beef cattle using 50K and imputed HD genotypes. *J. Anim. Sci.* 94, 1342–1353. doi:10.2527/JAS.2015-0126.

MacLeod, I. M., Bowman, P. J., Vander Jagt, C. J., Haile-Mariam, M., Kemper, K. E., Chamberlain, A. J., et al. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* 17, 144. doi:10.1186/S12864-016-2443-6.

Mao, F., Chen, L., Vinsky, M., Okine, E., Wang, Z., Basarab, J., et al. (2013). Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci.* 91, 2067–2076. doi:10.2527/jas.2012-5470.

McClellan, J., and King, M.-C. (2010). Genetic Heterogeneity in Human Disease. *Cell* 141, 210–217. doi:10.1016/J.CELL.2010.03.032.

Mehrban, H., Lee, D. H., Moradi, M. H., Ilcho, C., Naserkheil, M., and Ibáñez-Escriche, N. (2017). Predictive performance of genomic selection methods for carcass traits in Hanwoo beef cattle: Impacts of the genetic architecture. *Genet. Sel. Evol.* 49, 1. doi:10.1186/s12711-016-0283-0.

Melzer, N., Wittenburg, D., and Repsilber, D. (2013). Integrating Milk Metabolite Profile Information for the Prediction of Traditional Milk Traits Based on SNP Information for Holstein Cows. *PLoS One* 8, e70256. doi:10.1371/journal.pone.0070256.

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi:10.1093/GENETICS/157.4.1819.

Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* 6, 6–14. doi:10.2527/AF.2016-0002.

Montgomery, S. P., Sindt, J. J., Greenquist, M. A., Miller, W. F., Pike, J. N., Loe, E. R., et al. (2009). Plasma metabolites of receiving heifers and the relationship between apparent bovine respiratory disease, body weight gain, and carcass characteristics. *J. Anim. Sci.* 87, 328–333. doi:10.2527/jas.2008-0969.

Morgavi, D. P., Kelly, W. J., Janssen, P. H., and Attwood, G. T. (2013). “Rumen microbial (meta)genomics and its application to ruminant production,” in *Animal* (Cambridge University Press), 184–201. doi:10.1017/S1751731112000419.

Muggli-Cockett, N. E., Cundiff, L. V., and Gregory, K. E. (1992). Genetic analysis of bovine respiratory disease in beef calves during the first year of life1. *J. Anim. Sci.* 70, 2013–2019. doi:10.2527/1992.7072013x.

Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Grant, J. R., Mah, J., Wang, Z., et al. (2011). Associations of marker panel scores with feed intake and efficiency traits in beef cattle using preselected single nucleotide polymorphisms. *J. Anim. Sci.* 89, 3362–3371. doi:10.2527/JAS.2010-3362.

Mukiibi, R. (2019). Evaluation of genomic prediction accuracy of feed efficiency and the related traits with integration of differentially expressed gene SNPs.

Mukiibi, R., Vinsky, M., Keogh, K. A., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2018a). Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. *Sci. Rep.* 8, 7303. doi:10.1038/s41598-018-25605-3.

Mukiibi, R., Vinsky, M., Keogh, K. A., Stothard, P., Waters, S. M., and Li, C. (2018b). Transcriptome analysis of liver tissues in Charolais steers with divergent marbling phenotypes. in *Proceedings of the World Congress on Genetics Applied to Livestock Production* (Auckland, New Zealand), 11.626.

Neibergs, H. L., Seabury, C. M., Wojtowicz, A. J., Wang, Z., Scraggs, E., Kiser, J. N., et al. (2014). Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. *BMC Genomics* 15, 1164. doi:10.1186/1471-2164-15-1164.

Nkrumah, J. D., Okine, E. K., Mathison, G. W., Schmid, K., Li, C., Basarab, J. A., et al. (2006). Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J. Anim. Sci.* 84, 145–153. doi:10.2527/2006.841145x.

Preston, G. A., and Weinberger, D. R. (2005). Intermediate phenotypes in schizophrenia: a selective review. *Dialogues Clin. Neurosci.* 7, 165. doi:10.31887/DCNS.2005.7.2/GPRESTON.

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

Ramsey, R., Doye, D., Ward, C., McGrann, J., Falconer, L., and Bevers, S. (2005). Factors affecting beef cow-herd costs, production, and profits. *J. Agric. Appl. Econ.* 37, 91–99. doi:10.1017/s1074070800007124.

Rice, J. A., Carrasco-Medina, L., Hodgins, D. C., and Shewen, P. E. (2008). Mannheimia haemolytica and bovine respiratory disease. *Anim. Heal. Res. Rev.* 8, 117–128. doi:10.1017/S1466252307001375.

Sarup, P., Jensen, J., Ostersen, T., Henryon, M., and Sørensen, P. (2016). Increased prediction accuracy using a genomic feature model including prior information on quantitative trait locus regions in purebred Danish Duroc pigs. *BMC Genet.* 17, 11. doi:10.1186/s12863-015-0322-9.

Schlosser, P., Li, Y., Sekula, P., Raffler, J., Grundner-Culemann, F., Pietzner, M., et al. (2020). Genetic studies of urinary metabolites illuminate mechanisms of detoxification and

excretion in humans. *Nat. Genet.* 52, 167–176. doi:10.1038/s41588-019-0567-8.

Schröder, U. J., and Staufenbiel, R. (2006). Invited Review: Methods to Determine Body Fat Reserves in the Dairy Cow with Special Regard to Ultrasonographic Measurement of Backfat Thickness. *J. Dairy Sci.* 89, 1–14. doi:10.3168/JDS.S0022-0302(06)72064-1.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole blood transcriptomic analysis of beef cattle at arrival identifies potential predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. *PLoS One* 15, e0227507. doi:10.1371/JOURNAL.PONE.0227507.

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., et al. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18, 386. doi:10.1186/s12864-017-3754-y.

Selevsek, N., Chang, C. Y., Gillet, L. C., Navarro, P., Bernhardt, O. M., Reiter, L., et al. (2015). Reproducible and Consistent Quantification of the Saccharomyces cerevisiae Proteome by SWATH-mass spectrometry. *Mol. Cell. Proteomics* 14, 739–749. doi:10.1074/MCP.M113.035550.

Shin, S.-Y., Fauman, E. B., Petersen, A.-K., Krumsiek, J., Santos, R., Huang, J., et al. (2014). An atlas of genetic influences on human blood metabolites. *Nat. Genet.* 46, 543–550. doi:10.1038/ng.2982.

Silva, R. M. O., Fragomeni, B. O., Lourenco, D. A. L., Magalhães, A. F. B., Irano, N., Carvalheiro, R., et al. (2016). Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. *J. Anim. Sci.* 94, 3613–3623. doi:10.2527/JAS.2016-0401.

Smith, R. A. (2000). Effects of feedlot disease on economics, production and carcass value. in *The American Association of Bovine Practitioners Proceedings*, 125–128.

Snelling, W. M., Cushman, R. A., Fortes, M. R. S., Reverter, A., Bennett, G. L., Keele, J. W., et al. (2012). PHYSIOLOGY AND ENDOCRINOLOGY SYMPOSIUM: How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. *J. Anim. Sci.* 90, 1152–1165. doi:10.2527/JAS.2011-4581.

Snelling, W. M., Cushman, R. A., Keele, J. W., Maltecca, C., Thomas, M. G., Fortes, M. R. S., et al. (2013). BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection. *J. Anim. Sci.* 91, 537–552. doi:10.2527/JAS.2012-5784.

Snowder, G. (2009). Genetics, environment and bovine respiratory disease. *Anim. Health Res. Rev.* 10, 117–119. doi:10.1017/S1466252309990144.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2005). Influence of breed, heterozygosity, and disease incidence on estimates of variance components of respiratory disease in preweaned beef calves. *J. Anim. Sci.* 83, 1247–1261. doi:10.2527/2005.8361247x.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2006). Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *J. Anim. Sci.* 84, 1999–2008. doi:10.2527/jas.2006-046.

Srikanth, K., Lee, S.-H., Chung, K.-Y., Park, J.-E., Jang, G.-W., Park, M.-R., et al. (2020). A Gene-Set Enrichment and Protein–Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. *Genes (Basel).* 11, 316. doi:10.3390/GENES11030316.

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., et al. (2018). Genomic atlas of the human plasma proteome. *Nature* 558, 73–79. doi:10.1038/s41586-018-0175-2.

Sun, H. Z., Srithayakumar, V., Jiminez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics* 112, 3968–3977. doi:10.1016/J.YGENO.2020.07.014.

Theurer, M. E., Larson, R. L., and White, B. J. (2015). Systematic review and meta-analysis of the effectiveness of commercially available vaccines against bovine herpesvirus, Bovine viral diarrhea virus, Bovine respiratory syncytial virus, And parainfluenza type 3 virus for mitigation of bovine respiratory di. *J. Am. Vet. Med. Assoc.* 246, 126–142. doi:10.2460/javma.246.1.126.

Tizioto, P. C., Coutinho, L. L., Decker, J. E., Schnabel, R. D., Rosa, K. O., Oliveira, P. S. N., et al. (2015a). Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. *BMC Genomics* 16, 242. doi:10.1186/s12864-015-1464-x.

Tizioto, P. C., Kim, J. W., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015b). Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: An RNA-sequence analysis of the bronchial lymph node transcriptome. *PLoS One* 10, e0131459. doi:10.1371/journal.pone.0131459.

Torres-Vázquez, J. A., van der Werf, J. H. J., and Clark, S. A. (2018). Genetic and phenotypic associations of feed efficiency with growth and carcass traits in Australian Angus cattle. *J. Anim. Sci.* 96, 4521–4531. doi:10.1093/jas/sky325.

USDA (2011). Part IV: Health and Health Management on US Feedlots with a Capacity of 1,000 or More Head. Available at: https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_PartIV.pdf.

Wang, X., Zhang, Y., Zhang, X., Wang, D., Jin, G., Li, B., et al. (2017). The comprehensive liver transcriptome of two cattle breeds with different intramuscular fat content. *Biochem. Biophys. Res. Commun.* 490, 1018–1025. doi:10.1016/j.bbrc.2017.06.157.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Widmann, P., Reverter, A., Fortes, M. R. S., Weikard, R., Suhre, K., Hammon, H., et al. (2013). A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* 14, 798. doi:10.1186/1471-2164-14-798.

Wu, P., Wang, K., Zhou, J., Chen, D., Jiang, A., Jiang, Y., et al. (2021). A combined GWAS

approach reveals key loci for socially-affected traits in Yorkshire pigs. *Commun. Biol.* 4, 891. doi:10.1038/s42003-021-02416-3.

Xu, P., Liu, P., Zhou, C., Shi, Y., Wu, Q., Yang, Y., et al. (2019). A Multi-Omics Study of Chicken Infected by Nephropathogenic Infectious Bronchitis Virus. *Viruses* 11, 1070. doi:10.3390/V11111070.

Ye, S., Li, J., and Zhang, Z. (2020). Multi-omics-data-assisted genomic feature markers preselection improves the accuracy of genomic prediction. *J. Anim. Sci. Biotechnol.* 11, 109. doi:10.1186/S40104-020-00515-5.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (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, 36. doi:10.1186/s12864-019-6362-1.

Zhang, H., Yin, L., Wang, M., Yuan, X., and Liu, X. (2019). Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Front. Genet.* 10, 189. doi:10.3389/fgene.2019.00189.

Zhang, Y., Wang, Y., Li, Y., Wu, J., Wang, X., Bian, C., et al. (2020c). Genome-wide association study reveals the genetic determinism of growth traits in a Gushi-Anka F2 chicken population. *Heredity (Edinb).* 126, 293–307. doi:10.1038/s41437-020-00365-x.

Zhao, S., Jiang, H., Liang, Z.-H., and Ju, H. (2020). Integrating Multi-Omics Data to Identify Novel Disease Genes and Single-Neucleotide Polymorphisms. *Front. Genet.* 0, 1336. doi:10.3389/FGENE.2019.01336.

Zhao, W., Rasheed, A., Tikkanen, E., Lee, J.-J., Butterworth, A. S., Howson, J. M., et al. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat. Genet.* 49, 1450–1457. doi:10.1038/NG.3943.

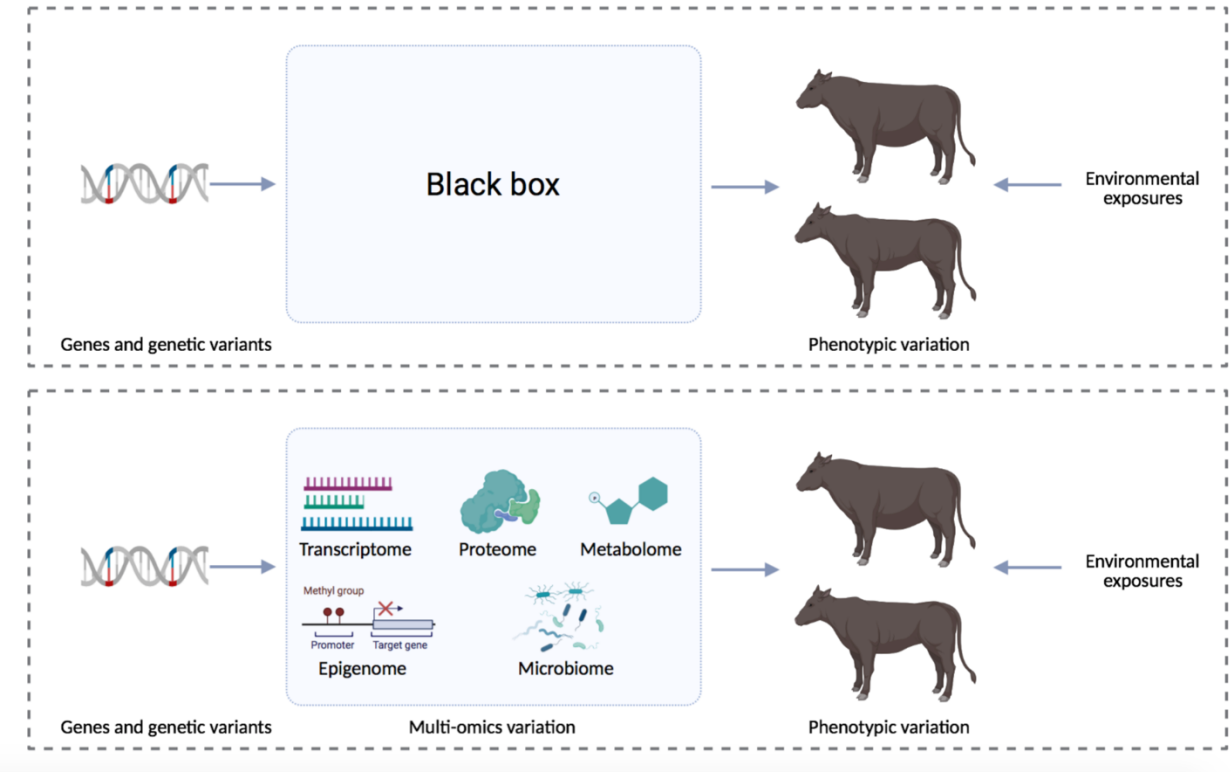

Figure 1.1 Schematic diagram of the intermediate phenotype lying between genome and phenome

Chapter 2. Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle

2.1 Abstract

Metabolites, substrates or products of metabolic processes, are involved in many biological functions, such as energy metabolism, signaling, stimulatory and inhibitory effects on enzymes and immunological defense. Metabolomic phenotypes are influenced by combination of genetic and environmental effects allowing for metabolome-genome-wide association studies (mGWAS) as a powerful tool to investigate the relationship between these phenotypes and genetic variants. The objectives of this study were to estimate genomic heritability and perform mGWAS and in silico functional enrichment analyses for a set of plasma metabolites in Canadian crossbred beef cattle.

Thirty-three plasma metabolites and 45,266 single nucleotide polymorphisms (SNPs) were available for 475 animals. Genomic heritability for all metabolites was estimated using genomic best linear unbiased prediction (GBLUP) including genomic breed composition as covariates in the model. A single-step GBLUP implemented in BLUPF90 programs was used to determine SNP *P* values and the proportion of genetic variance explained by SNP windows containing 10 consecutive SNPs. The top 10 SNP windows that explained the largest genetic variation for each metabolite were identified and mapped to detect corresponding candidate genes. Functional enrichment analyses were performed on metabolites and their candidate genes using the Ingenuity Pathway Analysis software.

Eleven metabolites showed low to moderate heritability that ranged from 0.09 ± 0.15 to 0.36 ± 0.15, while heritability estimates for 22 metabolites were zero or negligible. This result indicates that while variations in 11 metabolites were due to genetic variants, the majority are

largely influenced by environment. Three significant SNP associations were detected for betaine (rs109862186), L-alanine (rs81117935) and L-lactic acid (rs42009425) based on Bonferroni correction for multiple testing (family-wise error rate < 0.05). The SNP rs81117935 was found to be located within the *Catenin Alpha 2* gene (*CTNNA2*) showing a possible association with the regulation of L-alanine concentration. Other candidate genes were identified based on additive genetic variance explained by SNP windows of 10 consecutive SNPs. The observed heritability estimates and the candidate genes and networks identified in this study will serve as baseline information for research into the utilization of plasma metabolites for genetic improvement of crossbred beef cattle.

2.2 Introduction

The metabolic phenotype (or "metabotype") is a characteristic metabolite profile that depends on the interactions between genetic and environmental effects. Commonly, the metabolic phenotype of an individual is measured from easily accessible biofluids such as urine or blood (Nicholson and Lindon, 2008). Additionally, blood metabolites have been shown to be powerful tools for the indication of the nutritional and health status of humans and animals. For example, in humans, several blood metabolites have been identified as biomarkers for diseases (López-López et al., 2018). In livestock species, associations between metabolites and economically important traits such as feed efficiency (Karisa et al., 2014), growth performance (Widmann et al., 2013) and animal health (Montgomery et al., 2009) have been reported.

Metabolome-genome-wide association study is a powerful tool for identifying genetic variants underlying metabolic phenotypes and provides new opportunities to decipher the genetic basis of metabolic phenotypes. Importantly, metabolome-genome-wide association studies detect genetic variants that are functionally associated with metabolic phenotype variation. For example,

previous studies have reported that SNPs in the *glutamine synthase 2* gene (*GLS2*) were associated with glutamine in human serum, which provides a potential biological association, as the enzyme GLS2 catalyzes the hydrolysis of glutamine (Suhre et al., 2011; Kettunen et al., 2012). Furthermore, genome-wide hits with unknown gene function offer an opportunity to infer novel biological mechanisms of the SNP-metabolite association. For instance, Suhre et al. (2011) experimentally studied the association of the SNP rs7094971 inside the *solute carrier family 16, member 9* gene (*SLC16A9*) with carnitine and validated that the hitherto uncharacterized protein was indeed a carnitine transporter in *Xenopus* oocytes. Additionally, as metabolites lie between genomic and external phenotypes, they could explain the variation of external phenotypes by revealing biological mechanisms underlying the associations between them. Recent application of GWAS have successfully uncovered genetic variants that contribute to variation in both the external phenotype (e.g. type 2 diabetes) and the metabolic phenotype (e.g. fasting glucose levels) (Stranger et al., 2011).

Due to the rapidly growing number of candidate biomarkers and the increasing role of metabolites in genetic studies, the knowledge of the genetic basis of metabolites is therefore a prerequisite to evaluate new biomarkers and dissect the genetic architecture of metabolites. Until now, however, knowledge regarding the genetic level of metabolites in beef cattle has been limited. Thus, the objectives of this study were to estimate genomic heritability of 33 plasma metabolites in crossbred beef cattle, to identify genetic variants, genomic regions and candidate genes associated with metabolite variation, and to understand the biological functions and gene networks linked to these associations.

2.3 Material and Methods

2.3.1 Animal, blood samples and nuclear magnetic resonance (NMR) spectroscopy

All management and procedures involving live animals, where applicable, conformed to the guidelines outlined by the Canadian Council on Animal Care (2009); otherwise, existing data sets from the various Canadian research herds were used.

The dataset used in this study was obtained from the Phenomic Gap Project (McKeown et al., 2013). This project started in 2008 aiming to generate feed efficiency, carcass and meat quality phenotypes as well as genomic information for Canadian crossbred beef animals as previously described by Akanno et al. (2014). A total of 475 Canadian multibreed composite and crossbred beef cattle was used in this study. The animals comprised of bulls, slaughter steers, slaughter heifers and replacement heifers submitted to a feedlot feeding test from 2009 to 2012 and the test groups were labeled as contemporary groups. The population structure consisted of Beefbooster composite breed (n = 224) which is predominantly Charolais-based with infusion of Holstein, Maine Anjou, and Chianina (http://www.beefbooster.com), Hereford-Angus (n = 181) crosses, Charolais (n = 68) and Angus (n = 2).

Blood samples were collected in EDTA tubes from each animal by jugular venipuncture on the first day of the feedlot feeding test and immediately frozen at -80°C which is considered appropriate for storage. Our assumption is that all samples were affected equally by the freezing process if at all. Therefore, although the metabolite profiles may not be the same as those obtained from fresh samples, the freezing process should not be a source of variation for this study since all samples were frozen the same way according to best practice. Frozen blood samples were sent to the Metabolomics Innovation Center at University of Alberta, AB, Canada in 2014 for analysis. The variation in time of sample collection is expected to be captured under the ‘contemporary

group' variable applied in subsequent statistical analysis. Each frozen sample was thawed at room temperature then centrifuged at 10,000 rpm for 10 minutes to separate the plasma then filtered through 3kDa molecular weight cut-off filters (Merck Millipore Ltd., Darmstadt, Germany) to remove macromolecules, including lipids and proteins. As the filter tube manufacturer treats the filter membranes with glycerol as a preservative, filters were washed and centrifuged five times before use. Samples made up of less than 570 µl after filtration were diluted with HPLC water to ensure adequate volume for NMR acquisition. 5 mm NMR tube (New Era Enterprises Inc., NJ, USA) contained a total of 700 µl of total volume of 570 µl filtered serum, 60 µl DSS and 70 µl D2O. This mixture was vortexed and centrifuged shortly before it was transferred to an NMR tube for data acquisition. All metabolite concentrations obtained were adjusted by appropriate factors to account for the above dilutions, and represent the contents of the filtered samples, not the contents of the NMR tube.

Spectra were acquired on a 500MHz VNMRS spectrometer equipped with a 5mm cold probe (Agilent Technologies, CA, USA). The pulse sequence used was a 1D-noesy with a 990ms presaturation on water and a 4s acquisition period. Spectra were collected with 256 transients and 4 steady-state scans at 298K.

Spectra were zero filled to 64k points and Fourier transformed. Spectral phasing was performed on the spectra along with baseline correction. In total, 33 metabolites were identified and quantified with a targeted profiling approach using the Profiler and Library Manager modules in the same software which contains a total of 304 metabolites. Each spectrum was peer reviewed by a separate analyst and a final review pass was done on all of the spectra before exporting concentration results. Concentration measurements were adjusted to report metabolite concentrations after the filtration of the samples.

2.3.2 Genotyping, quality control and prediction of genomic breed composition

Animals were genotyped using Illumina BovineSNP50 v2 BeadChip (Illumina Inc., CA, USA) at Delta Genomics, Edmonton, AB, Canada. The genotypes were coded as 0, 1, and 2 and quality control was performed using the Synbreed package (Wimmer et al., 2012) in R statistical software. All markers on sex chromosomes and autosomal markers with minor allele frequency < 1%, call rate < 90%, and severe departure from Hardy-Weinberg equilibrium ($P < 10^{-5}$) were removed. Missing genotypes were imputed using Synbreed package. After quality control, 45,266 SNPs on 29 bovine autosomes for 475 individuals remained and were used for this study.

Genomic breed composition was predicted for all individuals using ADMIXTURE software (Alexander et al., 2009). To predict breed composition for each animal, a 10-fold cross-validation procedure was performed to find the best possible number of ancestors or breeds (K value). The value of $K = 4$ was chosen because it had the smallest cross-validation error and yielded the most accurate breed composition prediction based on prior knowledge. The four postulated ancestral breeds were Hereford, Angus, Charolais and Beefbooster TX line. The distribution of predicted genomic breed composition is shown in Figure 2.1. Estimates of genomic breed composition were fitted as covariates in the various statistical models to correct for population stratification and breed effects.

2.3.3 Phenotypic quality control

Phenotypic records included 33 plasma metabolite concentrations quantified from blood samples of 475 animals. A linear regression model implemented in R statistical software was used to assess the significance of all systematic effects associated with variation in plasma metabolites. Fixed factors found to be significant ($P < 0.05$) included contemporary groups (herd and birth year), animal type (bulls, slaughter steers, slaughter heifers and replacement heifers) and genomic

breed composition. These factors were subsequently included in the mixed model used for estimating heritability and GWAS. Contemporary group and animal type were fitted in the model as fixed class effect whereas breed fractions were fitted as fixed covariates. Residual values of the linear regression model were checked and those residuals with more or less than 3 standard deviations from the mean of residuals were considered as outliers and the associated records were excluded. The distribution of residuals after excluding outliers was close to a normal distribution (i.e. a bell shape without a big tail). The summary statistics of all metabolites after phenotypic quality control are given in Table 2.1. In general, the concentration of plasma metabolites ranged from 20.72 μM (L-methionine) to 5024.04 μM (L-lactic acid), on average.

2.3.4 Variance components and heritability estimation

Variance components and heritability of 33 metabolites were estimated using a single-trait animal model and genomic relationship matrix. The genomic relationship matrix was constructed based on total filtered SNP markers (i.e. 45,266 SNPs) and using one of VanRaden's formulations $ZZ' / 2\sum p_i(1-p_i)$, where Z contains centered genotypes codes and p_i is the minor allele frequency for locus *i* (VanRaden, 2008). The following mixed effect model (1) implemented in ASReml version 4.1 (Gilmour et al., 2015) was applied:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Wa} + \boldsymbol{e} \qquad (1)$$

Where $\boldsymbol{y}$ is a vector of phenotypic observation; $\boldsymbol{X}$ is the design matrix that relates the fixed effects to the observation and $\boldsymbol{b}$ is a vector of fixed effects of contemporary groups, animal type and genomic breed composition. $\boldsymbol{W}$ is a design matrix relating observations to random animal genetic effects; $\boldsymbol{a}$ is a vector of random additive polygenic effects that is assumed to be normally distributed as: $\boldsymbol{a} \sim N(\boldsymbol{0}, \boldsymbol{G}\sigma_{\boldsymbol{a}}^2)$, where $\boldsymbol{G}$ is genomic relationship matrix and $\sigma_{\boldsymbol{a}}^2$ is the additive

genetic variance, $\boldsymbol{e}$ is a vector of random residual effects that is assumed to be normally distributed as $\boldsymbol{e} \sim N(\mathbf{0}, \boldsymbol{I}\sigma_{\boldsymbol{e}}^2)$, with $\boldsymbol{I}$ being an identity matrix and $\sigma_{\boldsymbol{e}}^2$ is the residual error variance.

2.3.5 Metabolome-genome-wide association study

The genomic heritability obtained from model (1) was used to screen all metabolites for metabolome genome wide association analyses. Metabolites with zero or near zero heritability were excluded from mGWAS. Here, the SNP *P* values for 11 metabolites with non-zero heritability were determined using a single-step genomic BLUP (ssGBLUP) approach as described by Aguilar et al. (2019) and followed by the estimation of the proportion of additive variance explained by 10 consecutive SNP windows using a Weighted ssGBLUP (WssGBLUP) approach (Wang et al., 2012). Two iterations were applied in WssGBLUP. Both approaches were implemented in the BLUPF90 programs (Misztal et al., 2002). The mGWAS model used was similar to model (1) above except that $\boldsymbol{a}$ was assumed to follow $N(\mathbf{0}, \boldsymbol{H}\sigma_{\boldsymbol{a}}^2)$, where $\boldsymbol{H}$ is the matrix that combines genomic and pedigree information (Aguilar et al., 2010). The inverse of $\boldsymbol{H}$ for mixed model equations is:

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

$\boldsymbol{A}$ is the pedigree-based numerator relationship matrix for all animals, $\boldsymbol{A_{22}}$ is the numerator relationship matrix for genotyped animals, and matrix $\boldsymbol{G}$ is the genomic relationship matrix, where $\boldsymbol{G}$ was weighted as described by Wang et al. (2012) for the WssGBLUP method.

A rejection threshold based on Bonferroni correction for multiple testing (0.05/45266) was applied, which is equal to 5.96 in the -log10 scale. The quantile-quantile (Q-Q) plots of *P* values for each SNP were used to compare observed distributions of -log (*P* value) to the expected distribution under the null hypothesis for each metabolite. Manhattan plots of *P* values for each

SNP were also used to illustrate significant associations at the level of each chromosome for the metabolites. All plots were completed using the R package qqman (Turner, 2014).

2.3.6 Candidate gene identification

To identify a candidate gene, the surrounding region of each significant SNP was surveyed by expanding 100-kbp upstream and downstream, respectively. The 200-kbp region was defined based on the average linkage disequilibrium (r^2) between pairs of syntenic SNPs within this distance which is known to be 0.20 in a related beef cattle population (Lu et al., 2012).

Further, additional candidate genes associated with the top 10 SNP windows that explained the largest proportion of genetic variance for each metabolite from the WssGBLUP approach were identified. Positional candidate genes within 200-kbp regions and those inside the top 10 SNP windows were mapped on *Bos taurus* genome view in Biomart available at the Ensembl database UMD 3.1 version (Zerbino et al., 2018). The functions of all identified genes were manually searched from the literature to see if they had a previously identified relationship with the associated metabolites under investigation.

2.3.7 Analysis of least square means for significant SNPs

The least square mean of SNPs significantly associated with metabolites were assessed based on model (2) and implemented in R where applicable, to see how different allele combinations for these SNPs resulted in observed differences in the metabolite concentration.

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{SNP} + \boldsymbol{e} \qquad (2)$$

Where $\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{b}$ and $\boldsymbol{e}$ are the same as in model (1) and (2); $\boldsymbol{SNP}$ is a vector of genotype class 0, 1 and 2 fitted as a classification factor.

2.3.8 Functional enrichment analyses

The interpretation of mGWAS using metabolite concentrations as the target phenotype is a complicated task because their concentrations are influenced indirectly by mRNA and protein expression as well as directly by several environmental effects. Pathway analysis using prior knowledge improves the interpretation of mGWAS data and provides insight from the genetics of biochemical conversions and biological functions. Functional analyses for the genes associated with each metabolite were performed using Ingenuity Pathway Analysis software (IPA; www.Ingenuity.com). Several lists including metabolites (PubChem CID) and candidate genes (Bovine Entrez gene IDs) in Table 2.S1 were imported in IPA for biological function analysis and network construction. Biological functions were considered significantly enriched if the *P* value for the overlap comparison test between the input list and the knowledge base of IPA for a given biological function was less than 0.05. Identification of significant pathways in biological processes was described in detail by Calvano et al. (2005). The analysis was performed following IPA default setting and parameters were set to allow the network to show indirect relationships for the imported metabolite and gene lists. Indirect relationships assist in the identification of other metabolites/genes that were not among the ones in the input list but may be associated with them based on the IPA biological reference. In addition, the resulting gene networks are scored and then sorted based on the score not based on *P* value, as multiple testing for this sort of analysis is not feasible.

2.4 Results

2.4.1 Heritability estimates

Eleven metabolites showed low to moderate heritability that ranged from 0.09 ± 0.15 (succinic acid) to 0.36 ± 0.15 (choline), while heritability estimates for 22 metabolites were zero or negligible. Table 2.2 shows the results of all metabolites with heritability.

2.4.2 SNP association, candidate genes and genetic effects

Three significant SNP associations were detected for betaine (rs109862186), L-alanine (rs81117935) and L-lactic acid (rs42009425) based on Bonferroni correction for multiple testing (family-wise error rate < 0.05) (Table 2.3, Figures 2.2 – 2.4). The SNPs were located on chromosome 5, 11 and 22, respectively. The SNP rs81117935 was found within the *catenin alpha 2* gene (*CTNNA2*), while the other two SNPs were not mapped to any known candidate gene (Table 2.4).

In addition to the identified significant SNPs, the WssGBLUP method also identified more genomic regions associated with heritable metabolites based on additive genetic variance explained by SNP windows of 10 consecutive SNPs. The proportion of additive genetic variance explained by top 10 SNP windows and genes mapped in these windows are shown in Table 2.S1. The SNP window (107,403,824 - 107,704,991 bp) located on chromosome 5 was found to be associated with citric acid and explained the highest proportion of additive genetic variance (4.21%) while the SNP window (35,619,632 - 36,428,58 bp) with the lowest proportion of additive genetic variance (0.62%) was located on chromosome 26 and associated with L-lactic acid. A total of 368 unique genes were identified within the selected SNP windows (Table 2.S1). Further, five SNP windows showed pleiotropic effects on two or more metabolites and were mapped to 17 candidate genes (Table 2.5).

The least square means of the genotypic classes are given in Figure 2.5. All three significant SNPs (rs109862186, rs81117935 and rs42009425) showed characteristics of additivity with the associated metabolite as concentration either increased or decreased with the number of "B" alleles for the three genotypic classes.

2.4.3 Functional enrichment analyses

The eleven heritable metabolites and their candidate genes were significantly enriched ($P < 0.05$) for biological functions related to cellular, tissue, and organ development, cell-to-cell signaling and interaction, molecular transport, small molecule biochemistry, lipid metabolism, carbohydrate metabolism, and cellular growth and proliferation. Additionally, one of the most informative networks (Figure 2.6) was related to lipid metabolism and cell-to-cell signaling and interaction with betaine and some of its candidate genes.

2.5 Discussion

2.5.1 Heritability estimates

Metabolites have the potential to serve as biomarkers for production traits and diseases in livestock (Montgomery et al., 2009), and the concentration of biomarkers should not vary too much over the short term within a healthy individual because such variation could undermine the predictive association in a single sample (Nicholson et al., 2011b). Most highly conserved metabolites are also highly heritable (Yousri et al., 2014) and less influenced by the environmental changes. In this study, we performed a baseline investigation into the heritability of plasma metabolites in crossbred beef cattle and identified potential associations between heritable metabolites and SNP markers. As certain metabolites are essential for growth and health, knowledge of the genetic parameters of these important metabolites could trigger directional

selection towards regulating their concentration in metabolic processes. For instance, alanine is an essential amino acid for T cell activation (Ron-Harel et al., 2019) which affects immunity level. Here, a total of eleven metabolites out of thirty-three showed low to moderate heritability, suggesting their potential as biomarkers for genetic selection. Betaine and choline which showed moderate heritability in this study have been previously identified to be associated with residual feed intake in beef cattle (Karisa et al., 2014), thus, they could potentially be used as biomarkers for improving feed efficiency in beef cattle. The majority of the metabolites with negligible heritability may be largely influenced by environmental effects such as age, gender, nutrition, medication, and possibly underlying diseases (Beuchel et al., 2019). The non-heritable status of these metabolites may be used as a guide to animal management. For example, ruminants fed silage-based diets are likely to ingest ethanol because of ethanol production in fermented feeds (Nishino and Shinde, 2007) and the process of ethanol detoxification in liver could affect splanchnic nutrient metabolism (Obitsu et al., 2013). Ethanol showed a negligible heritability in this study, which suggests that the variation of ethanol concentration may be mainly affected by management factors such as feed.

In a related study that utilized milk metabolites from dairy cattle, Buitenhuis et al. (2013) found heritability estimates that were similar to estimates observed for five metabolites from the current study. Although, these studies are not completely comparable, this finding corroborates the possible existence of a genetic basis for plasma metabolites. In addition, the negligible heritability or large standard error observed for some of the metabolites may be due to the limited number of animals utilized. Thus, further study may be warranted as this is the first attempt to characterise the genetic basis of plasma metabolites in crossbred beef cattle.

2.5.2 SNP association, candidate genes and genetic effects

Genetic profiling of plasma metabolites has been previously studied in other species to assess their value as biomarkers for disease prediction (López-López et al., 2018). Recently, metabolomics GWAS was performed to identify genomic regions associated with variation in milk metabolites in dairy cattle (Buitenhuis et al., 2013). To the best of our knowledge, this study is the first attempt at profiling the genetic basis of plasma metabolites in crossbred beef cattle. The SNPs and candidate genes identified here revealed the potential association between metabolomics and genetics, which could help fill the knowledge gap that exist between genetic level and external phenotype. The possible signals detected in this study were associated with betaine, L-alanine and L-lactic acid and the peaks for significant additive SNPs including rs109862186, rs81117935 and rs42009425 were on chromosome 5, 11, and 22. Two of the SNPs rs109862186 and rs42009425 showed no evidence of a candidate gene within 200-kbp distance, however, SNP rs42009425 associated with L-lactic acid was reported to be associated with clinical mastitis in French Holstein cattle (Marete et al., 2018). The SNP rs81117935 associated with L-alanine was found to be located within the candidate gene *CTNNA2* which is one of three human alpha-catenin genes. Alpha-catenin functions as a linking protein between cadherins and actin-containing filaments of the cytoskeleton (Cooper and Hausman, 2000), however, it is not known whether *CTNNA2* gene may regulate the concentration of L-alanine in bovine blood. The least square mean results (Figure 2.5) showed that the concentration of L-alanine was significantly ($P < 0.05$) greater in individuals that are homozygotes for the "A" allele of SNP rs81117935 while no significant differences existed for the other two genotypic classes. Our finding suggests that *CTNNA2* gene may play a role in the regulation of plasma L-alanine which requires further investigation.

Further, several candidate genes associated with heritable metabolites were mapped inside the selected SNP windows of 10 consecutive SNPs based on WssGBLUP analyses. Here, *choline kinase alpha* gene (*CHKA*) which is associated with choline was mapped inside the SNP window (46,143,465 – 46,796,930 bp) on chromosome 29. This gene encodes an enzyme called choline kinase alpha (Hosaka et al., 1992) which catalyzes the phosphorylation of choline to phosphocholine (Aoyama et al., 2004) as a first step in the biosynthesis pathway of phosphatidylcholine (Lacal, 2001). Phosphatidylcholine is one of the most abundant phospholipids in all mammalian cell membranes (van der Veen et al., 2017) and plays a critical role in membrane structure and also in cell signaling (Lacal, 2001). The importance of phospholipid metabolism in regulating lipid, lipoprotein and whole-body energy metabolism has been reviewed by van der Veen et al. (2017). Lipid metabolism has been previously identified as an important biological function in relation to beef cattle residual feed intake (Chen et al., 2011; Alexandre et al., 2015; Mukiibi et al., 2018). Therefore, the relationship between *CHKA* gene and choline metabolite used in this study has potential value for improving feed efficiency in beef cattle. Interestingly, several overlapped SNP windows were also identified, which indicates that either two metabolites were controlled by the same gene or by different genes within a SNP window (Table 2.5). The substantial polygenic and pleiotropic nature of the metabolite variation observed in the current study have been previously reported in human metabolomics studies (Hu et al., 2018; Gallois et al., 2019).

Several reasons may lead to the few significant SNPs identified. Firstly, variation in metabolite concentrations may be due to the polygenic nature of the genes underlying the variation. Polygenic inheritance for primary metabolites have been reported in humans (Tanha et al., 2021) and plants (Rowe et al., 2008; Chan et al., 2010; Wen et al., 2014) and could potentially exist in

beef cattle as evident in our study that utilized primary metabolites. Secondly, the crossbred nature of our studied population could lead to inconsistent linkage disequilibrium across multiple populations (De Roos et al., 2009). Thirdly, the ability to identify SNPs and quantitative trait loci with large effects on any of the metabolites depends partly on the amount of variation in metabolite concentration that can be attributed to genetic source. Here, low to moderate heritability were observed for some of the metabolites studied. Marker density is another factor that may lead to identification of fewer significant SNPs associated with variation in metabolites. In this study, 50K SNP panel was used for mGWAS, however, most causative SNPs may not be included in this panel and thus, would likely not be detected. Studies involving other beef cattle traits have shown that increasing marker density from 50K to 7.8 million SNPs can capture more additive genetic variance and can detect additional or novel significant SNPs (Wang et al., 2020; Zhang et al., 2020). Therefore, high-density SNP marker panel or whole-genome sequence data are suggested for future studies. Lastly, a stringent significance threshold based on Bonferroni correction for multiple testing was imposed to identify significant SNPs and exclude false positive results. However, compared with traditional GWAS, metabolites are highly correlated to other similar metabolites and often cannot be considered as independent. The traditional multiple testing methods may therefore eliminate some valuable SNPs. Some groups have computed the Bonferroni correction by counting all the metabolites (Gieger et al., 2008; Illig et al., 2010; Suhre et al., 2011), while a few other groups have adopted a less stringent strategy by taking into account the number of independent metabolites as determined by a principal component analysis to adjust for multiple test correction (Demirkan et al., 2012).

2.5.3 Functional enrichment analyses

A one-to-one metabolite-to-gene correspondence is not known *a priori* (Nicholson et al., 2011a) but functional enrichment analyses could provide enriched functions and networks of metabolites and identified candidate genes to give a whole picture of gene-metabolite associations. Some biological functions that are significantly enriched may help us improve understanding of molecular factors for some important traits, such as feed efficiency. The eight most significantly enriched biological functions for beef cattle feed efficiency included lipid metabolism, amino acid metabolism, carbohydrate metabolism, energy production, molecular transport, small molecule biochemistry, cellular development, and cell death and survival (Cantalapiedra-Hijar et al., 2018). Our results supplement the part played by genetic and molecular factors for these functions, thus, available data with both information (i.e., metabolite data and feed efficiency related traits) could be used to elucidate this hypothesis. Detailed insight into the specific pathways that are affected by variation in metabolites is a useful first step to select the most likely hypotheses. A good example is betaine which is widely distributed within the animal body (Xia et al., 2018) and was reported to enhance the synthesis of methylated compounds such as phospholipids as well as directly influence lipid metabolism (Huang et al., 2008). In addition, a recent study showed that insulin was associated with phospholipid alterations, but the mechanism is still not clear (Chang et al., 2019). Interestingly, the enriched pathway constructed by IPA showed a relationship between betaine, insulin and phospholipids and provides new insight into the connection between them (Figure 2.6), however, this connection requires experimental validation.

2.6 Conclusion

This study estimated heritability of 33 plasma metabolites for crossbred beef cattle and found low to moderate heritability for 11 metabolites, which provides evidence for the genetic

basis underlying the variation of metabolite concentrations. Three significant SNP associations were detected for betaine (rs109862186), L-alanine (rs81117935) and L-lactic acid (rs42009425) which suggest that the genetic effects may be largely polygenic. The SNP rs81117935 was found to be within *CTNNA2* gene which is possibly associated with the regulation of L-alanine concentration in bovine blood. Other candidate genes were identified based on additive genetic variance explained by SNP windows of 10 consecutive SNPs. The observed heritability estimates and candidate genes and networks identified in this study will serve as baseline information for further research into the utilization of plasma metabolites for genetic improvement of crossbred beef cattle.

2.7 References

Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., and Misztal, I. (2019). Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol.* 51, 28. doi:10.1186/s12711-019-0469-3.

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

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle - An independent validation study. *J. Anim. Sci.* 92, 2896–2904. doi:10.2527/jas.2013-7374.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109.

Alexandre, P. A., Kogelman, L. J. A., Santana, M. H. A., Passarelli, D., Pulz, L. H., Fantinato-Neto, P., et al. (2015). Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16, 1073. doi:10.1186/s12864-015-2292-8.

Aoyama, C., Liao, H., and Ishidate, K. (2004). Structure and function of choline kinase isoforms in mammalian cells. *Prog. Lipid Res.* 43, 266–281. doi:10.1016/j.plipres.2003.12.001.

Beuchel, C., Becker, S., Dittrich, J., Kirsten, H., Toenjes, A., Stumvoll, M., et al. (2019). Clinical and lifestyle related factors influencing whole blood metabolite levels – A comparative analysis of three large cohorts. *Mol. Metab.* 29, 76–85. doi:10.1016/j.molmet.2019.08.010.

Buitenhuis, A. J. J., Sundekilde, U. K. K., Poulsen, N. A. A., Bertram, H. C. C., Larsen, L. B. B., and Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *J. Dairy Sci.* 96, 3285–3295. doi:10.3168/jds.2012-5914.

Calvano, S. E., Xiao, W., Richards, D. R., Felciano, R. M., Baker, H. V., Cho, R. J., et al. (2005). A network-based analysis of systemic inflammation in humans. *Nature* 437, 1032–1037. doi:10.1038/nature03985.

Cantalapiedra-Hijar, G., Abo-Ismail, M., Carstens, G. E., Guan, L. L., Hegarty, R., Kenny, D. A., et al. (2018). Review: Biological determinants of between-animal variation in feed efficiency of growing beef cattle. *Animal* 12, S321–S335. doi:10.1017/S1751731118001489.

Chan, E. K. F., Rowe, H. C., Hansen, B. G., and Kliebenstein, D. J. (2010). The Complex Genetic Architecture of the Metabolome. *PLoS Genet.* 6, e1001198. doi:10.1371/journal.pgen.1001198.

Chang, W., Hatch, G. M., Wang, Y., Yu, F., and Wang, M. (2019). The relationship between phospholipids and insulin resistance: From clinical to experimental studies. *J. Cell. Mol. Med.* 23, 702–710. doi:10.1111/jcmm.13984.

Chen, Y., Gondro, C., Quinn, K., Herd, R. M., Parnell, P. F., and Vanselow, B. (2011). Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim. Genet.* 42, 475–490. doi:10.1111/j.1365-2052.2011.02182.x.

Cooper, G. M., and Hausman, R. E. (2000). *The Cell: A Molecular Approach*. Washington, DC: ASM press.

De Roos, A. P. W., Hayes, B. J., and Goddard, M. E. (2009). Reliability of genomic predictions across multiple populations. *Genetics* 183, 1545–1553. doi:10.1534/genetics.109.104935.

Demirkan, A., van Duijn, C. M., Ugocsai, P., Isaacs, A., Pramstaller, P. P., Liebisch, G., et al. (2012). Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet.* 8, e1002490. doi:10.1371/journal.pgen.1002490.

Gallois, A., Mefford, J., Ko, A., Vaysse, A., Julienne, H., Ala-Korpela, M., et al. (2019). A comprehensive study of metabolite genetics reveals strong pleiotropy and heterogeneity across time and context. *Nat. Commun.* 10, 4788. doi:10.1038/s41467-019-12703-7.

Gieger, C., Geistlinger, L., Altmaier, E., Hrabé de Angelis, M., Kronenberg, F., Meitinger, T., et al. (2008). Genetics meets metabolomics: A genome-wide association study of metabolite profiles in human serum. *PLoS Genet.* 4, 1000282. doi:10.1371/journal.pgen.1000282.

Gilmour, A. R., Gogel, B. J., and Welham, S. J. (2015). *ASReml User Guide Functional Specification*. Hemel Hempstead: VSN International Ltd.

Hosaka, K., Tanaka, S., Nikawa, J. ichi, and Yamashita, S. (1992). Cloning of a human choline kinase cDNA by complementation of the yeast *cki* mutation. *FEBS Lett.* 304, 229–232. doi:10.1016/0014-5793(92)80625-Q.

Hu, Y., Tan, L.-J., Chen, X.-D., Liu, Z., Min, S.-S., Zeng, Q., et al. (2018). Identification of novel potentially pleiotropic variants associated with osteoporosis and opbesity using the

cFDR method. *J. Clin. Endocrinol. Metab.* 103, 125–138. doi:10.1210/jc.2017-01531.

Huang, Q. C., Xu, Z. R., Han, X. Y., and Li, W. F. (2008). Effect of dietary betaine supplementation on lipogenic enzyme activities and fatty acid synthase mRNA expression in finishing pigs. *Anim. Feed Sci. Technol.* 140, 365–375. doi:10.1016/j.anifeedsci.2007.03.007.

Illig, T., Gieger, C., Zhai, G., Römisch-Margl, W., Wang-Sattler, R., Prehn, C., et al. (2010). A genome-wide perspective of genetic variation in human metabolism. *Nat. Genet.* 42, 137–141. doi:10.1038/ng.507.

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., et al. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livest. Sci.* 165, 200–211. doi:10.1016/j.livsci.2014.03.002.

Kettunen, J., Tukiainen, T., Sarin, A. P., Ortega-Alonso, A., Tikkanen, E., Lyytikäinen, L. P., et al. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* 44, 269–276. doi:10.1038/ng.1073.

Lacal, J. C. (2001). Choline kinase: a novel target for antitumor drugs. *IDrugs* 4, 419–26. doi:10.1042/bj1800559.

López-López, Á., López-Gonzálvez, Á., Barker-Tejeda, T. C., and Barbas, C. (2018). A review of validated biomarkers obtained through metabolomics. *Expert Rev. Mol. Diagn.* 18, 557–575. doi:10.1080/14737159.2018.1481391.

Lu, D., Sargolzaei, M., Kelly, M., Li, C., Voort, G. Vander, Wang, Z., et al. (2012). Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Genet.* 3, 152. doi:10.3389/fgene.2012.00152.

Marete, A., Sahana, G., Fritz, S., Lefebvre, R., Barbat, A., Lund, M. S., et al. (2018). Genome-wide association study for milking speed in French Holstein cows. *J. Dairy Sci.* 101, 6205–6219. doi:10.3168/jds.2017-14067.

McKeown, L., Aalhus, J., Larsen, I., Stothard, P., and Wang, Z. (2013). Bridging the "Phenomic Gap": Creation of a database containing phenotypes and genotypes for economically important traits for beef cattle. Edmonton, Canada. Final Report to the Alberta Livestock and Meat Agency.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., and Lee, D. H. (2002). BLUPF90 and related programs (BGF90). in *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, 21–22.

Montgomery, S. P., Sindt, J. J., Greenquist, M. A., Miller, W. F., Pike, J. N., Loe, E. R., et al. (2009). Plasma metabolites of receiving heifers and the relationship between apparent bovine respiratory disease, body weight gain, and carcass characteristics. *J. Anim. Sci.* 87, 328–333. doi:10.2527/jas.2008-0969.

Mukiibi, R., Vinsky, M., Keogh, K. A., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2018). Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. *Sci. Rep.* 8, 7303. doi:10.1038/s41598-018-25605-3.

Nicholson, G., Rantalainen, M., Li, J. V., Maher, A. D., Malmodin, D., Ahmadi, K. R., et al. (2011a). A genome-wide metabolic QTL analysis in europeans implicates two Loci shaped

by recent positive selection. *PLoS Genet.* 7, e1002270. doi:10.1371/journal.pgen.1002270.

Nicholson, G., Rantalainen, M., Maher, A. D., Li, J. V., Malmodin, D., Ahmadi, K. R., et al. (2011b). Human metabolic profiles are stably controlled by genetic and environmental variation. *Mol. Syst. Biol.* 7, 525. doi:10.1038/msb.2011.57.

Nicholson, J. K., and Lindon, J. C. (2008). Systems biology: Metabonomics. *Nature* 455, 1054–1056. doi:10.1038/4551054a.

Nishino, N., and Shinde, S. (2007). Ethanol and 2,3-butanediol production in whole-crop rice silage. *Grassl. Sci.* 53, 196–198. doi:10.1111/j.1744-697x.2007.00089.x.

Obitsu, T., Nishimura, K., Udaka, Y., Sugino, T., and Taniguchi, K. (2013). "Effects of ethanol on splanchnic nutrient metabolism in sheep at different intake levels," in *Energy and protein metabolism and nutrition in sustainable animal production* (Wageningen Academic Publishers), 441–442. doi:10.3920/978-90-8686-781-3_164.

Ron-Harel, N., Ghergurovich, J. M., Notarangelo, G., LaFleur, M. W., Tsubosaka, Y., Sharpe, A. H., et al. (2019). T Cell Activation Depends on Extracellular Alanine. *Cell Rep.* 28, 3011-3021.e4. doi:10.1016/j.celrep.2019.08.034.

Rowe, H. C., Hansen, B. G., Halkier, B. A., and Kliebenstein, D. J. (2008). Biochemical networks and epistasis shape the *Arabidopsis thaliana* metabolome. *Plant Cell* 20, 1199–1216. doi:10.1105/tpc.108.058131.

Stranger, B. E., Stahl, E. A., and Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 187, 367–383. doi:10.1534/genetics.110.120907.

Suhre, K., Shin, S. Y., Petersen, A. K., Mohney, R. P., Meredith, D., Wägele, B., et al. (2011). Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 477, 54–62. doi:10.1038/nature10354.

Tanha, H. M., Sathyanarayanan, A., and Nyholt, D. R. (2021). Genetic overlap and causality between blood metabolites and migraine. *Am. J. Hum. Genet.* 108, 2086–2098. doi:10.1016/J.AJHG.2021.09.011.

Turner, S. D. (2014). qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*, 005165. doi:10.1101/005165.

van der Veen, J. N., Kennelly, J. P., Wan, S., Vance, J. E., Vance, D. E., and Jacobs, R. L. (2017). The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim. Biophys. Acta - Biomembr.* 1859, 1558–1572. doi:10.1016/j.bbamem.2017.04.006.

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

Wang, H., Misztal, I., Aguilar, I., Legarra, A., and Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res. (Camb).* 94, 73–83. doi:10.1017/S0016672312000274.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies

of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Wen, W., Li, D., Li, X., Gao, Y., Li, W., Li, H., et al. (2014). Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun.* 5, 3438. doi:10.1038/ncomms4438.

Widmann, P., Reverter, A., Fortes, M. R. S., Weikard, R., Suhre, K., Hammon, H., et al. (2013). A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* 14, 798. doi:10.1186/1471-2164-14-798.

Wimmer, V., Albrecht, T., Auinger, H. J., and Schön, C. C. (2012). Synbreed: A framework for the analysis of genomic prediction data using R. *Bioinformatics* 28, 2086–2087. doi:10.1093/bioinformatics/bts335.

Xia, Y., Chen, S., Zhu, G., Huang, R., Yin, Y., and Ren, W. (2018). Betaine inhibits interleukin-1β production and release: Potential mechanisms. *Front. Immunol.* 9, 2670. doi:10.3389/fimmu.2018.02670.

Yousri, N. A., Kastenmüller, G., Gieger, C., Shin, S.-Y., Erte, I., Menni, C., et al. (2014). Long term conservation of human metabolic phenotypes and link to heritability. *Metabolomics* 10, 1005–1017. doi:10.1007/s11306-014-0629-y.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi:10.1093/nar/gkx1098.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (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, 36. doi:10.1186/s12864-019-6362-1.

Table 2.1 Descriptive statistics for 33 plasma metabolites: number of animals per metabolite (n), mean, standard deviation (SD), coefficient of variation (CV), minimum (Min.) and maximum (Max.). Unit: µM

Trait	n	Mean	SD	CV	Min.	Max.
1-methylhistidine	435	56.26	22.71	0.40	15.34	136.31
2-hydroxybutyrate	460	41.23	17.02	0.41	12.26	94.48
acetic acid	462	264.60	256.05	0.97	33.40	2,056.21
betaine	448	111.67	52.97	0.47	29.62	298.33
creatine	451	127.59	44.39	0.35	41.98	262.67
citric acid	448	120.27	65.38	0.54	15.61	338.45
choline	456	346.37	173.98	0.50	61.35	960.08
ethanol	404	61.38	84.91	1.38	13.53	560.94
D-glucose	452	837.40	692.11	0.83	68.42	3,731.80
glycine	451	378.65	162.32	0.43	90.38	896.70
glycerol	452	511.10	354.71	0.69	15.68	1,532.64
fumaric acid	300	23.85	8.48	0.36	10.75	66.11
formic acid	454	30.34	28.25	0.93	9.46	370.87
L-tyrosine	475	65.51	19.32	0.29	22.88	119.90
L-phenylalanine	454	67.54	19.54	0.29	27.53	125.61
L-alanine	446	390.34	148.99	0.38	104.46	852.47
L-proline	465	129.58	41.02	0.32	42.09	257.82
L-isoleucine	465	52.85	19.88	0.38	15.11	120.63
L-histidine	450	76.09	28.57	0.38	23.35	150.45
lysine	460	70.34	26.19	0.37	15.24	154.49
L-lactic acid	450	5,024.04	2,790.01	0.56	885.17	15,976.05
pyruvic acid	321	87.56	81.42	0.93	14.23	395.75
succinic acid	448	58.47	34.46	0.59	14.86	280.58
3-hydroxybutyric acid	457	86.65	41.66	0.48	18.29	272.70
creatinine	451	132.14	57.85	0.44	30.77	308.61
L-glutamine	441	58.97	23.00	0.39	14.35	119.97
L-leucine	475	93.08	39.48	0.42	25.63	302.17
L-methionine	193	20.72	4.49	0.22	12.08	33.77
3-hydroxyisovaleric acid	155	32.38	13.02	0.40	11.70	79.06
L-valine	454	147.16	49.58	0.34	49.88	313.97
acetone	260	35.97	19.84	0.55	12.47	125.08
methanol	447	135.47	76.28	0.56	31.35	383.19
dimethyl sulfone	449	46.86	19.41	0.41	15.31	128.60

Table 2.2 Estimates of additive variance (σ_a^2), residual variance (σ_e^2), heritability (h^2) and their standard error (SE) for eleven plasma metabolites[1].

Trait	σ_a^2	σ_e^2	h^2	SE
choline	6,598.90	11,545.80	0.36	0.15
creatinine	1,051.67	1,947.73	0.35	0.17
betaine	402.10	783.09	0.34	0.16
pyruvic acid	1,027.32	2,007.84	0.34	0.24
L-lactic acid	639,240	2,268,490	0.22	0.16
citric acid	477.13	1,719.37	0.22	0.15
creatine	160.55	843.99	0.16	0.15
D-glucose	17,497.10	100,579.00	0.15	0.14
acetone	29.39	185.01	0.14	0.21
L-alanine	768.05	7,824.22	0.09	0.13
succinic acid	78.47	838.28	0.09	0.15

[1]Metabolites with zero or near zero heritability estimates were not listed.

Table 2.3 SNPs significantly associated with metabolites: chromosome (Chr), position of SNP on chromosome (bp), minor allele and minor allele frequency (MAF), nucleotide of SNP, *P* values of significant test and Bonferroni correction of *P* values.

Trait	SNP	Chr	Position (bp)	Minor allele and MAF	Nucleotide (major/minor allele)	*P*	Bonferroni correction
betaine	rs109862186	5	118,820,845	B (0.18)	T/C	7.63E-07	0.03
L-alanine	rs81117935	11	54,765,154	A (0.45)	T/C	9.10E-07	0.04
L-lactic acid	rs42009425	22	41,109,447	A (0.19)	A/G	9.94E-07	0.04

Table 2.4 200-kpb regions around the significant SNPs: chromosome (Chr), position of the region on chromosome (bp), gene in the regions and the location of the gene compared to SNP location.

Trait	Chr	Position (bp)	Gene name	Gene location compared to SNP location
betaine	5	118,720,845 - 118,920,845	--	-
L-alanine	11	54,665,154 - 54,865,154	*CTNNA2*	SNP is within gene
L-lactic acid	22	41,009,447 - 41,209,447	-	-

Table 2.5 Chromosome (Chr), position of overlapped windows (bp) and genes in overlap windows.

Traits	Chr	Position (bp)	Gene name
acetone, L-lactic acid	1	28,675,718 - 29,049,389	*GBE1*
L-alanine, choline	7	13,336,301 - 13,632,174	*IER2, STX10, TRMT1, LYL1, NACC1, NFIX, CACNA1A*
L-alanine, betaine	19	24,357,241 - 24,917,540	*RAP1GAP2, SPATA22, OR1G1, ASPA, TRPV1, TRPV3*
L-alanine, creatine	21	49,290,972 - 49,623,230	*GEMIN2, PNN*
creatine, choline	28	15,916,594 - 16,124,333	*ANK3*

Table 2.S1 Genes that were identified within the top 10 SNP windows for heritable metabolites: proportion of additive genetics variance explained by each region, chromosome (Chr), start and end of the region on chromosome, genes in the regions.

Trait	Variance explained by region	Chr	Start (bp)	End (bp)	Gene name
Acetone	2.45	1	28648176	29049389	*GBE1*
	2.06	9	98124370	98526918	*MAP3K4, AGPAT4, PRKN*
	1.33	10	44821159	45069536	*GNG2, RTRAF, NID2*
	1.12	5	63870367	64565816	*UHRF1BP1L, ACTR6, SCYL2*
	1.01	5	111313740	111790193	*SYNGR1, TAB1, MGAT3, MIEF1, ATF4, RPS19BP1, CACNA1I, ENTHD1*
	0.95	7	65379758	65689219	*NMUR2*
	0.95	5	57424855	58282365	*NABP2, RNF41, SMARCC2, MYL6, MYL6B, ESYT1, ZC3H10, PA2G4, ERBB3, RPS26, IKZF4, SUOX, RAB5B, CDK2, PMEL, DGKA, PYM1, MMP19, DNAJC14, ORMDL2, SARNP, GDF11, CD63, RDH5, BLOC1S1, ITGA7, METTL7B, OR10P1, OR2AP1*
	0.78	23	9617524	10026534	*FKBP5, ARMC12, CLPS, LHFPL5, SRPK1, SLC26A8, MAPK14*
	0.67	5	30275164	30912029	*TMBIM6, FMNL3, PRPF40B, FAM186B, MCRS1, KCNH3, SPATS2, DNAJC22, C1QL4, TROAP, PRPH, TUBA1C, TUBA1A, TUBA1B, LMBR1L, DHH*
	0.66	6	12942582	13295514	*CAMK2D*
Betaine	1.86	18	33602408	33891066	
	1.71	14	20952022	21452744	*SPIDR, H3F3C, PRKDC, UBE2V2*
	1.71	9	96215762	96715173	*TULP4, TMEM181, DYNLT1, SYTL3, EZR*
	1.59	9	97732365	98047172	*IGF2R, SLC22A1, SLC22A2, SLC22A3, PLG*
	1.48	2	96021111	96857370	*CREB1, METTL21A, CCNYL1, FZD5, PLEKHM3, CRYGD, CRYGC, CRYGB*
	1.17	19	24357241	24917540	*RAP1GAP2, OR1G1, SPATA22, ASPA, TRPV3, TRPV1*
	1.10	7	56766754	57175519	
	1.06	2	108665467	109020542	
	1.04	11	46307696	46701073	*NT5DC4, CKAP2L, IL1A, IL1B, IL37, IL36G, IL36A, IL36B, IL36RN, IL1F10, IL1RN*
	0.97	4	75741220	76123840	
Creatine	4.16	21	29563115	29923689	*PCSK6, SNRPA1*
	2.36	28	15812621	16124333	*ANK3*
	1.53	6	54369253	54625529	
	1.46	21	49242650	49623230	*SEC23A, GEMIN2, TRAPPC6B, PNN*

	1.22	26	27894025	28368749	*SORCS1*
	1.19	14	64767268	65300637	*NCALD, GRHL2*
	1.19	11	66073158	66554571	*C1D, WDR92*
	1.06	8	93786059	94115663	
	0.99	7	3069070	3388020	*WNT9A, PRSS38, SNAP47, JMJD4*
	0.99	3	69940346	70313793	*TYW3, CRYZ*
Citric acid	4.21	5	107403824	107704991	
	1.94	1	144411032	144758824	*SLC37A1, PDE9A, WDR4, NDUFV3, PKNOX1*
	1.88	5	7733715	8063248	
	1.62	3	96751550	97221503	
	1.62	6	75380980	76407633	
	1.37	21	51461130	52239094	*LRFN5*
	1.14	17	62311865	62623054	*GLRB, PDGFC, TBX3, TBX5*
	1.12	17	43052005	43459410	
	1.08	26	36905943	37336603	*GFRA1, CCDC172, PNLIPRP3, PNLIP*
	0.89	13	29277185	29647662	*FAM107B*
Choline	2.64	26	49366950	49736388	*EBF3*
	2.62	7	13336301	13632174	*CACNA1A, IER2, STX10, NACC1, TRMT1, LYL1, NFIX*
	1.56	1	1983902	2462297	*C1H21orf62, PAXBP1, SYNJ1, CFAP298, EVA1C, URB1*
	1.42	29	46143465	46796930	*ALDH3B2, UNC93B1, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, C29H11orf24, PPP6R3, GAL*
	1.35	28	15916594	16250507	*ANK3*
	1.06	12	12557821	12977211	*TNFSF11, FAM216B*
	0.99	2	69753620	70442711	*CCDC93, INSIG2*
	0.99	1	18099231	18578077	*TMPRSS15, CHODL*
	0.97	6	25033485	25556469	*PPP3CA*
	0.96	19	63625695	64007021	*CACNG5, CACNG4, CACNG1, HELZ*
D-Glucose	3.72	13	25298145	25681327	*KIAA1217*
	1.83	13	23925250	24688809	*PIP4K2A, ARMC3, MSRB2, PTF1A, OTUD1*
	1.51	2	1039834	1316010	*CYFIP1, TUBGCP5, CCDC115, IMP4, PTPN18*
	1.14	24	29013292	29349786	*CDH2*
	1.09	6	64466274	64733603	
	0.96	13	75310875	75604222	*TNNC2, ACOT8, ZSWIM3, ZSWIM1, SPATA25, NEURL2, CTSA, PLTP, PCIF1, ZNF335, MMP9, SLC12A5, NCOA5, CD40*
	0.93	1	61996164	62696719	
	0.92	1	132388576	132768483	

	0.85	6	41858119	42155077	*KCNIP4*
	0.78	9	45699093	46351157	*HACE1*
L-Alanine	4.06	11	54730416	55170733	*CTNNA2*
	2.87	21	49290972	49646269	*SEC23A, GEMIN2, TRAPPC6B, PNN*
	2.56	1	45957974	46402519	*IMPG2, SENP7, TRMT10C, PCNP, ZBTB11*
	1.99	7	62839580	63231799	*CSNK1A1, ARHGEF37, PPARGC1B, PDE6A*
	1.26	3	96751550	97221503	
	1.08	19	23220366	23605621	*PITPNA, SLC43A2, SCARF1, RILP, PRPF8, TLCD2, MIR22, WDR81, SERPINF2, SERPINF1, SMYD4, RPA1, RTN4RL1*
	1.08	19	24357241	24917540	*RAP1GAP2, OR1G1, SPATA22, ASPA, TRPV3, TRPV1*
	0.94	7	13336301	13632174	*CACNA1A, IER2, STX10, NACC1, TRMT1, LYL1, NFIX*
	0.92	7	86461228	86823742	*EDIL3*
	0.66	5	107302660	107590490	*TEAD4, TULP3, RHNO1, FOXM1, TEX52, NRIP2, ITFG2, FKBP4, DDX11, WASHC1, IQSEC3*
L-Lactic acid	1.83	1	28675718	29073969	*GBE1*
	0.90	7	6698584	8376994	*AP1M1, FAM32A, CIB3, HSH2D, RAB8A, TPM4, OR10H1*
	0.87	12	15154465	15573724	*NUFIP1, GPALPP1, GTF2F2, KCTD4, TPT1*
	0.85	20	8897859	9270217	*ZNF366, PTCD2, MRPS27*
	0.80	9	93552330	94154628	
	0.74	12	79456283	79814959	*STK24, SLC15A1, DOCK9*
	0.74	19	49247372	49627477	*MILR1, POLG2, DDX5, CEP95, SMURF2, KPNA2, BPTF*
	0.68	13	29821617	30337948	*SUV39H2, DCLRE1C, MEIG1, OLAH, ACBD7, RPP38, NMT2, FAM171A1*
	0.63	22	43478537	43834473	*PXK, RPP14, ABHD6, DNASE1L3, FLNB*
	0.62	26	35619632	36428581	*TRUB1, ATRNL1*
Pyruvic acid	1.71	16	48544837	48879755	
	1.19	11	4525511	4918860	*REV1*
	1.06	12	61686477	62517886	
	1.02	20	52816141	53201978	
	0.97	2	110598183	111155237	*PAX3*
	0.96	24	2094260	2466721	*GALR1, ZNF516*
	0.95	8	110455623	110921940	*BRINP1*
	0.93	10	45323831	45687660	*PIF1, RBPMS2, OAZ2, ZNF609, TRIP4*
	0.83	6	15294694	15633977	
	0.79	24	2795809	3143836	
Succinic acid	3.91	3	106289280	106611961	*SMAP2, COL9A2, ZMPSTE24, TMCO2, RLF*

	2.79	27	15291371	15686768	*FAM149A, CYP4V2, KLKB1, F11, MTNR1A, FAT1*
	2.41	7	41321459	41742027	*MGAT1, ZFP62, BTNL9, TRIM7*
	2.40	1	19182174	19829143	
	1.45	7	56910268	57334500	*YIPF5, KCTD16*
	1.30	1	131261969	131642361	*PIK3CB, FAIM, CEP70*
	1.09	6	108026956	108340883	*SH3BP2, TNIP2, FAM193A, RNF4*
	1.04	2	129768904	130187033	*RPL11, ID3, E2F2, ASAP3, TCEA3, ZNF436, HNRNPR*
	0.89	11	24050822	24317047	
	0.87	19	44632448	45005031	*TMUB2, ATXN7L3, UBTF, SLC4A1, RUNDC3A, SLC25A39, GRN, FAM171A2, ITGA2B, GPATCH8, FZD2*
Creatinine	1.64	20	676757	1052840	
	1.59	9	101640789	102007389	*C9H6orf118, PDE10A*
	1.45	20	17617134	17986264	*ZSWIM6*
	1.32	22	52504436	52955533	*SMARCC1, CSPG5, ELP6, SCAP, PTPN23, NGP, KLHL18, KIF9*
	1.14	10	77151640	77518423	*PLEKHG3, SPTB, CHURC1, GPX2, RAB15, FNTB, MAX*
	1.02	27	43066841	43459156	
	0.98	1	73077982	73487973	*XXYLT1, FAM43A, LSG1*
	0.89	7	56766754	57175519	
	0.87	29	29194986	29595613	*PKNOX2, FEZ1, EI24, STT3A, CHEK1, ACRV1, SSLP1, PATE2*
	0.86	12	11955547	12513750	*VWA8, DGKH, AKAP11*

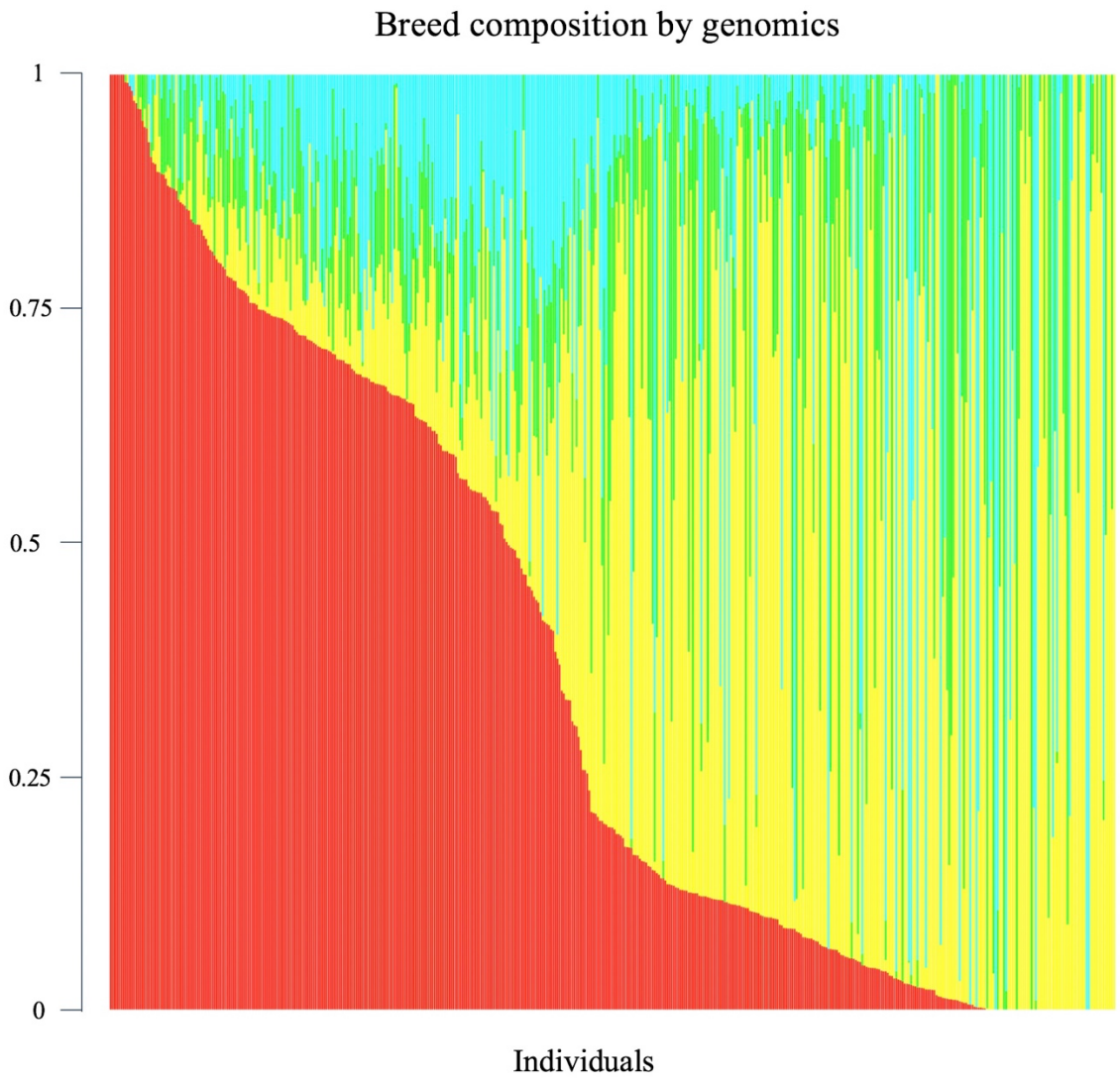

Figure 2.1 Distribution of predicted genomic breed composition of crossbred beef cattle population (n = 475). Beefbooster is red, Angus is yellow, Hereford is green, Charolais is blue.

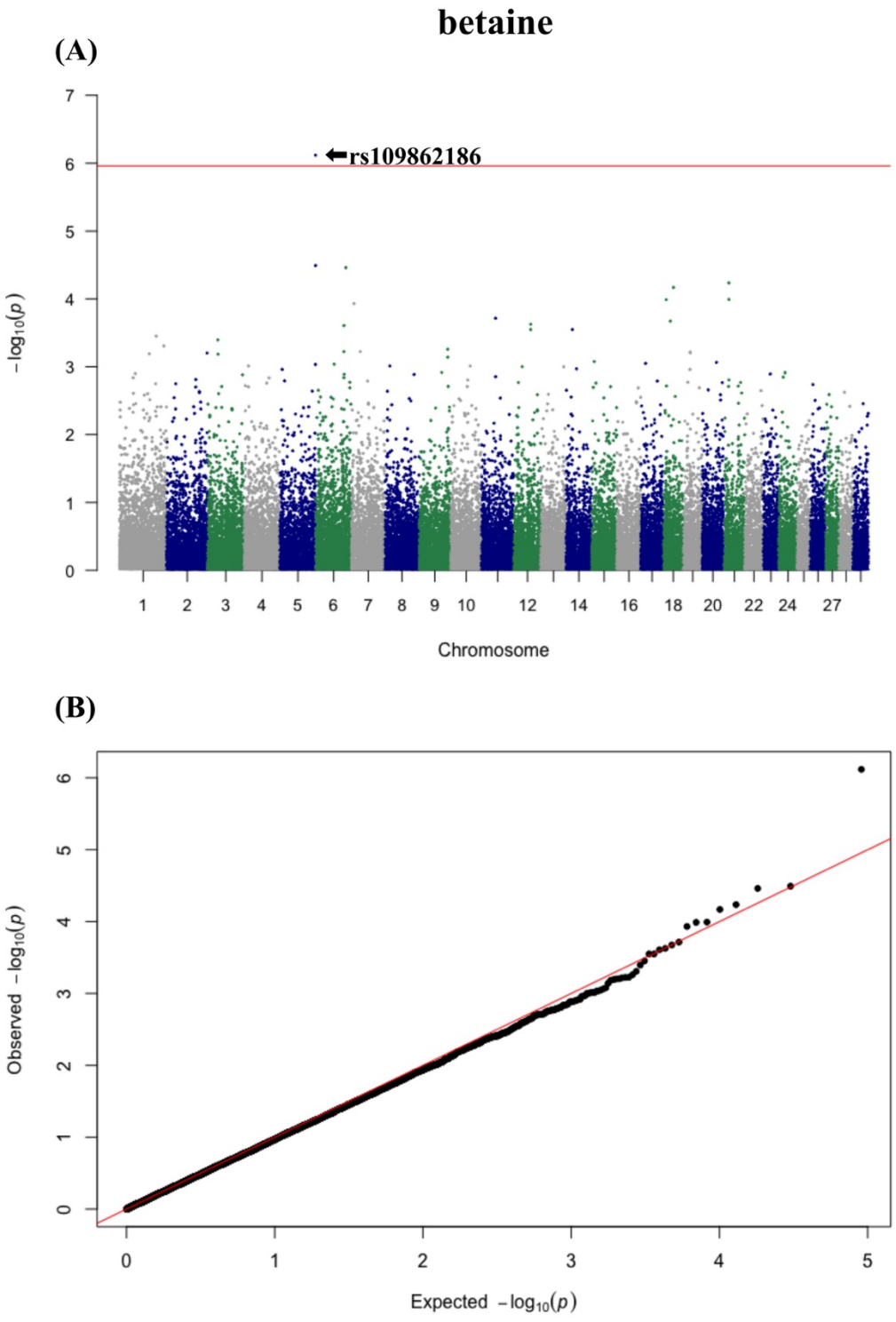

Figure 2.2 Manhattan plot (A) and QQ plot (B) for betaine, significant SNPs were determined by Bonferroni correction (red line).

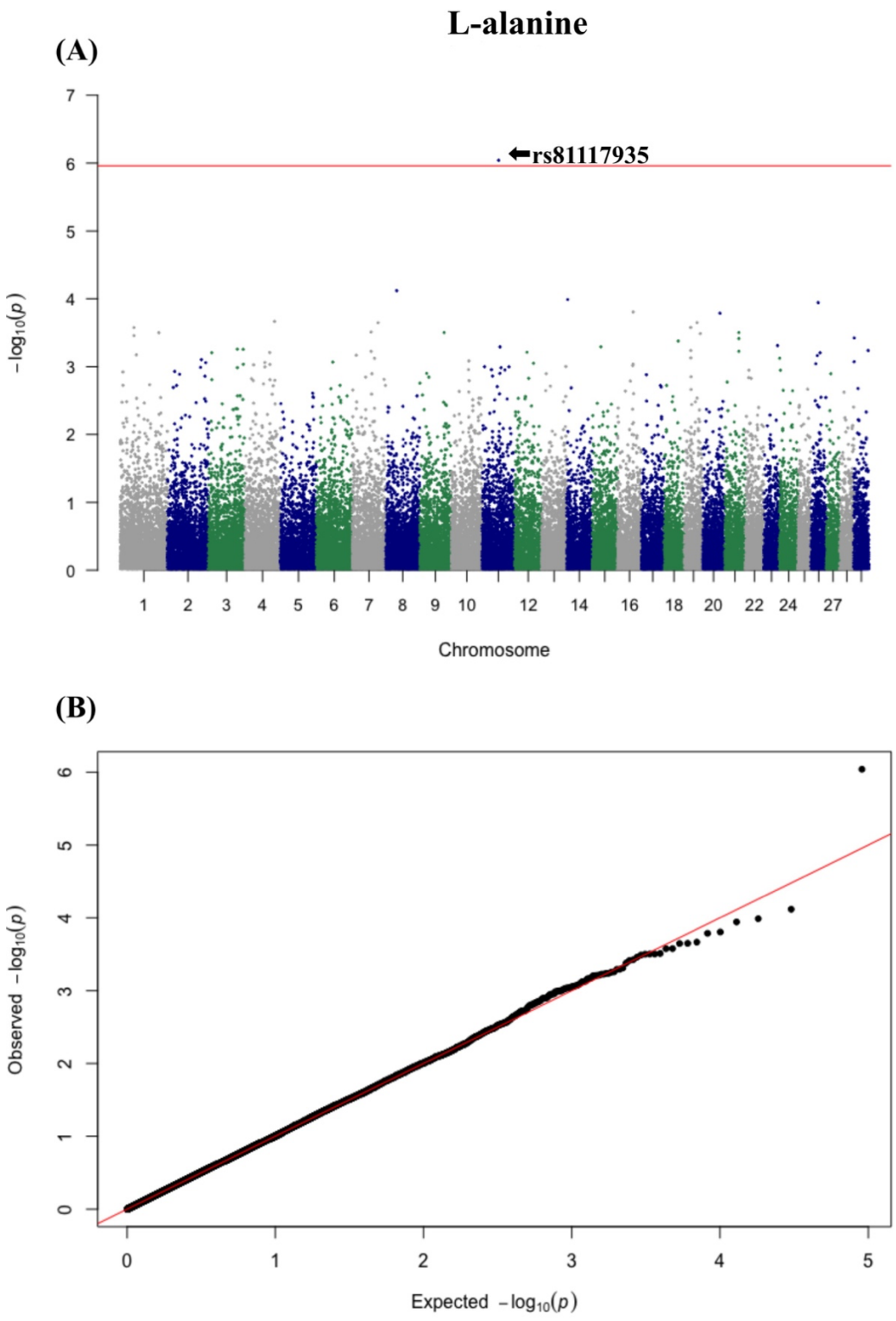

Figure 2.3 Manhattan plot (A) and QQ plot (B) for L-alanine, significant SNPs were determined by Bonferroni correction (red line).

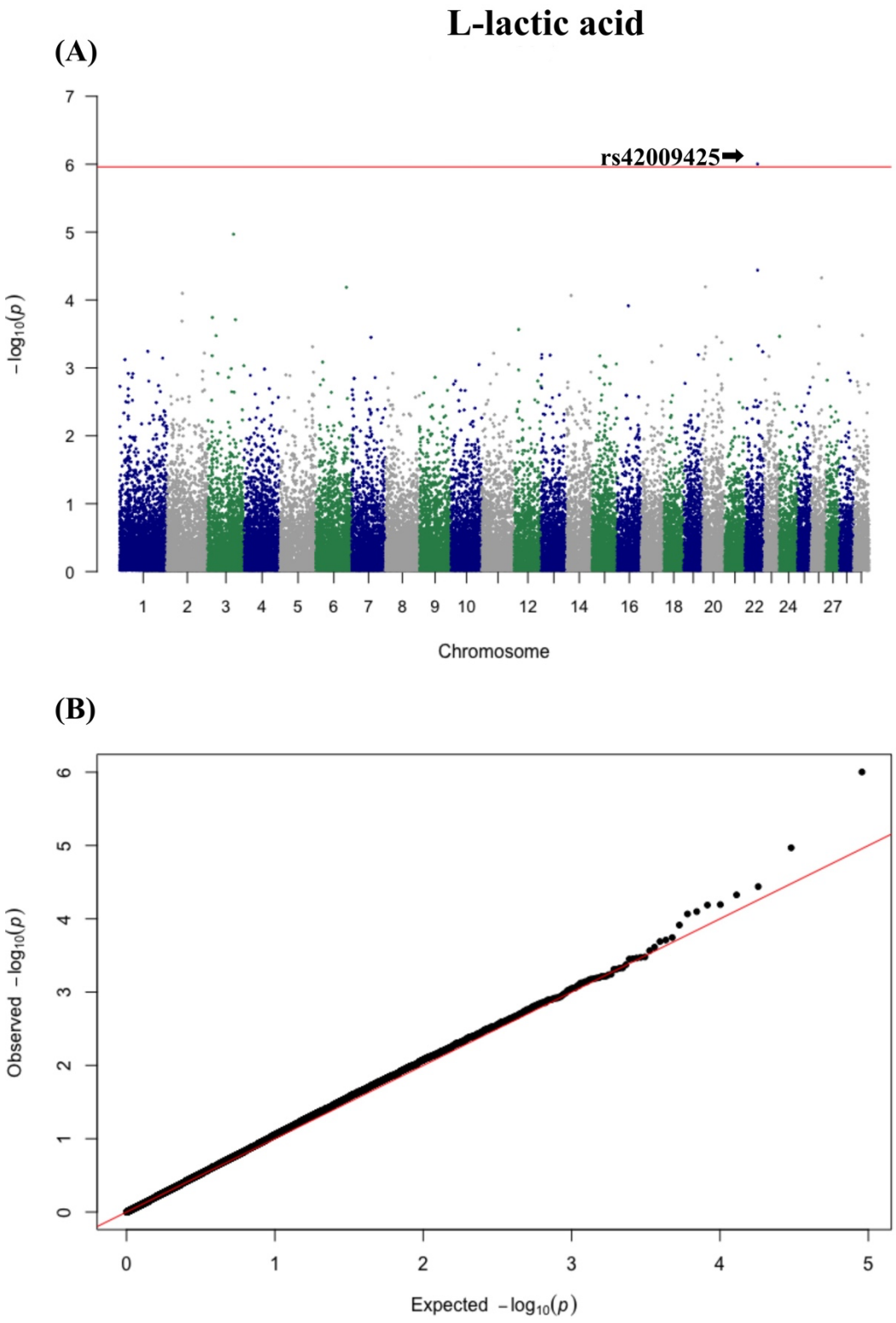

Figure 2.4 Manhattan plot (A) and QQ plot (B) for L-lactic acid, significant SNPs were determined by Bonferroni correction (red line).

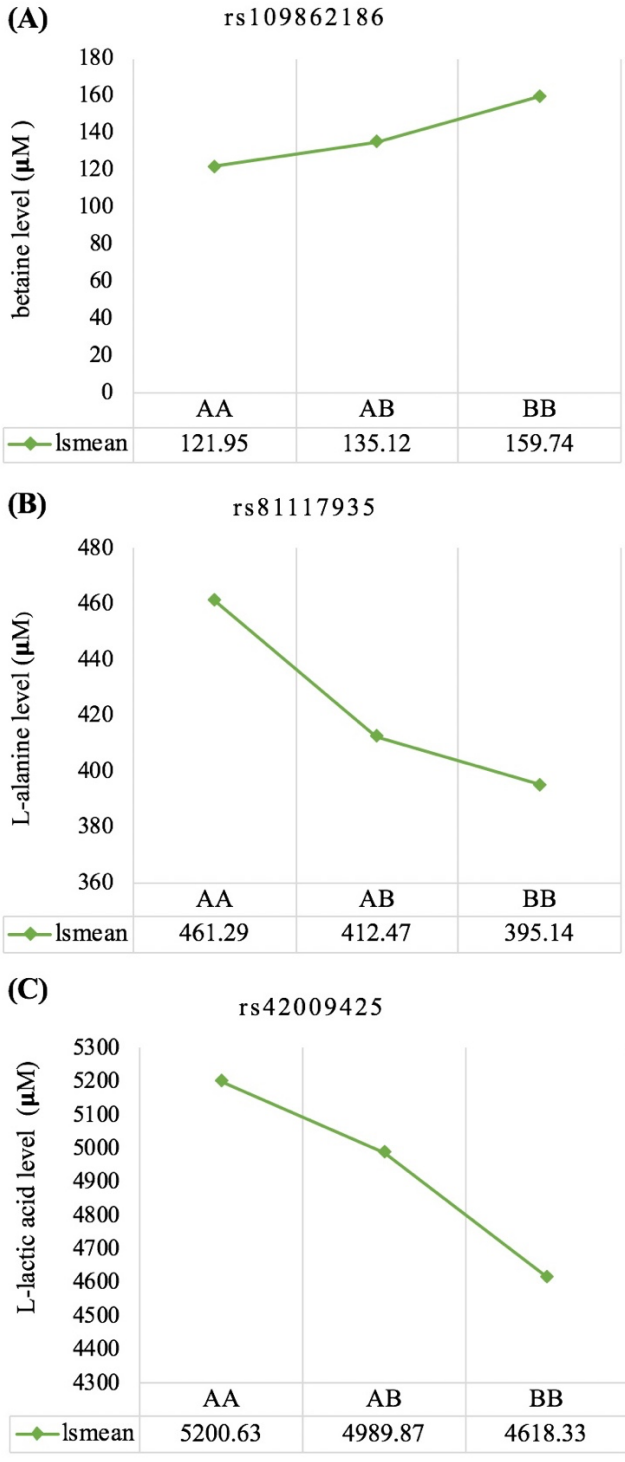

Figure 2.5 Least square means for the genotypic classes of significant SNPs associated with betaine (A), L-alanine (B) and L-lactic acid (C), respectively. All three significant SNPs (rs109862186, rs81117935 and rs42009425) showed characteristics of additivity with the associated metabolite.

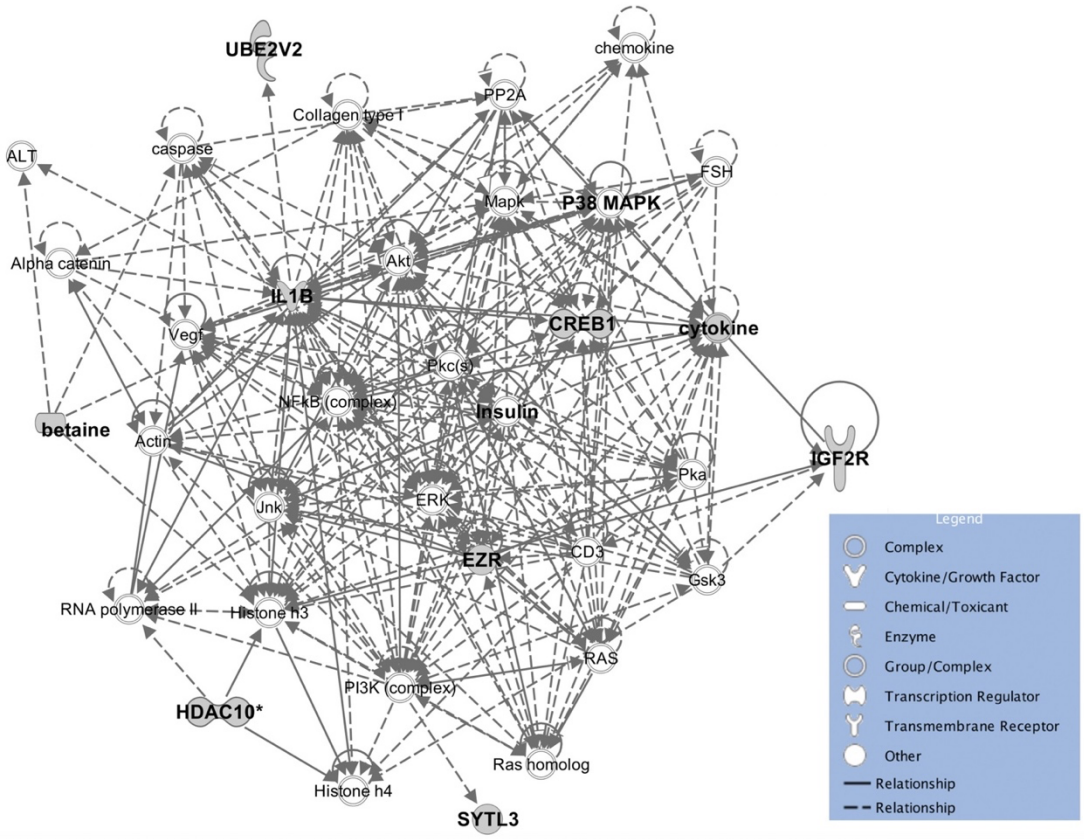

Figure 2.6 The enrichment network for betaine and associated genes, and the molecules in IPA database. The enriched pathway predicted by IPA showed a potential relationship between betaine, insulin and phospholipids.

Chapter 3. Identification of candidate genes and enriched biological functions for feed efficiency traits by integrating plasma metabolites and imputed whole genome sequence variants in beef cattle

3.1 Abstract

Feed efficiency is one of the key determinants of beef industry profitability and sustainability. However, the cellular and molecular background behind feed efficiency is largely unknown. This study combines imputed whole genome DNA variants and 31 plasma metabolites to dissect genes and biological functions/processes that are associated with residual feed intake (RFI) and its component traits including dry matter intake (DMI), average daily gain (ADG), and metabolic body weight (MWT) in beef cattle.

Regression analyses between feed efficiency traits and plasma metabolites in a population of 493 crossbred beef cattle identified 5 (L-valine, lysine, L-tyrosine, L-isoleucine, and L-leucine), 4 (lysine, L-lactic acid, L-tyrosine, and choline), 1 (citric acid), and 4 (L-glutamine, glycine, citric acid, and dimethyl sulfone) plasma metabolites associated with RFI, DMI, ADG, and MWT (*P*-value < 0.1), respectively. Combining the results of metabolome-genome wide association studies using 10,488,742 imputed SNPs, 40, 66, 15, and 40 unique candidate genes were identified as associated with RFI, DMI, ADG, and MWT (*P*-value < 1×10^{-5}), respectively. These candidate genes were found to be involved in some key metabolic processes including metabolism of lipids, molecular transportation, cellular function and maintenance, cell morphology and biochemistry of small molecules.

This study identified metabolites, candidate genes and enriched biological functions/processes associated with RFI and its component traits through the integrative analyses of metabolites with phenotypic traits and DNA variants. Our findings could enhance the

understanding of biochemical mechanisms of feed efficiency traits and could lead to improvement of genomic prediction accuracy via incorporating metabolite data.

3.2 Introduction

Feeding-related costs are the major expense in beef cattle enterprises, representing 55% - 75% of total production costs (Ramsey et al., 2005; Ahola and Hill, 2012; Nielsen et al., 2013). Reducing feed inputs per unit of production could significantly improve profitability by 9 to 33% in beef production (Archer et al., 2004). Additionally, with the projected increase of the global population to 9.6 billion by the year 2050, the growing demand for beef is likely to put more pressure on already limited production resources such as water, land, fertilizers and labor (Gerber et al., 2013). Moreover, studies have shown that more feed efficient beef cattle consume less feed for the same amount of beef produced, and meanwhile, have a reduced methane emission (Hegarty et al., 2007). Therefore, improvements in feed efficiency of beef cattle can increase producer profitability and simultaneously lower the environmental footprint of beef production.

Residual feed intake (RFI) is an important indicator of feed efficiency, which is usually defined as the difference between an animal's actual daily dry matter intake (DMI) and the expected daily DMI given the animal's average daily gain (ADG) and metabolic body weight (MWT) (Koch et al., 1963). Currently, measuring individual animal feed intake to calculate RFI is a complex and expensive process. Numerous studies in beef cattle have revealed moderate to high heritability estimates (0.16 - 0.68) for RFI (Herd and Bishop, 2000; Arthur et al., 2001; Nkrumah et al., 2006; Mao et al., 2013), and thus make RFI suitable for genetic/genomic selection of efficient beef cattle. Over the decades, genome-wide association studies (GWAS) have detected thousands of single nucleotide polymorphisms (SNPs) and hundreds of candidate genes associated with RFI in beef cattle (Bolormaa et al., 2011; Abo-Ismail et al., 2014; Seabury et al., 2017; Zhang

et al., 2020). However, cellular and molecular functions associated with transcriptomic, metabolomic and proteomic levels of omic data, and detailed knowledge regarding the biological processes involved in feed efficiency still remain largely unknown. Metabolites are substrates or products of metabolic processes and are the results of combined endogenous and exogenous production (Fontanesi, 2016), thus metabolites are considered as intermediate phenotypes between the genomic (base) and phenotypic (top) levels (Fontanesi, 2016). Integration of metabolomic data into feed efficiency studies could help reveal the relationship between animal genetics and physiological phenotypes (i.e., RFI and its component traits), thereby increasing the fundamental understanding of biological functions related to feed efficiency and improving genetic/genomic selection efficacy in beef cattle. Therefore, the objective of this study was to use metabolites as intermediate phenotypes to study genes and biological functions/processes related to feed efficiency in beef cattle. In this study, feed efficiency data were collected from a beef cattle population consisting of 493 crossbred bulls, heifers, and steers. Thirty-one metabolites and their concentration levels (μM) were quantified from plasma of these animals on the first day of feedlot tests. Linear regression models were applied to identify metabolites associated with RFI and its component traits (DMI, ADG, and MWT). Whole genome sequence variants were imputed and used in metabolome-genome wide association studies (mGWAS) to identify significant SNPs for trait associated metabolites. Candidate genes were mapped based on significant SNPs and gene functional enrichment analyses were subsequently performed on candidate genes of each trait to predict biological functions/processes associated with feed efficiency in beef cattle.

3.3 Material and Methods

3.3.1 Animal population, data collection of feed efficiency traits and metabolites

All animals in this study were cared for according to the guidelines of the Canadian Council on Animal Care (2009). The population of animals was obtained from the Phenomic Gap Project that aimed to generate phenotypes of feed efficiency, carcass and meat quality as well as genomic data for Canadian crossbred beef animals (McKeown et al., 2013). Details of animal management, the herd, and animal breeds were previously described (Akanno et al., 2014; Abo-Ismail et al., 2018; Li et al., 2020; Zhang et al., 2020). In summary, the population used in this study consisted of 493 crossbred bulls (n = 93), heifers (n = 125) and steers (n = 275) that were born between 2002 to 2011. These animals were from five different commercial herds and they were tested in feedlots from 2003 to 2012 (McKeown et al., 2013). The major breed components were primarily Charolais (n = 73), Hereford-Angus crosses (n = 191) and a Beefbooster composite breed (predominantly Charolais-based, n = 229). The GrowSafe system (GrowSafe Systems Ltd., Airdrie, Alberta, Canada) was used to measure the feed intake of finishing calves at the feeding test station for a period of 76 to 112 days. Serial body weights (BW) in kg were measured for each animal at the beginning and end of the test and at approximately 14-day intervals during the test. The daily DMI in kg was calculated as an average of dry matter intake over the test period and further standardized according to the energy content of the diet. The initial BW in kg at the start of the feeding test and the ADG in kg were derived from a linear regression of the serial BW measurements against time (day) (Nkrumah et al., 2007; Lu et al., 2013; Mao et al., 2013; Zhang et al., 2020). The MWT in kg was calculated as midpoint $BW^{0.75}$ while the midpoint BW was computed as the sum of the initial BW in kg and the product of ADG multiplied by half of the days on test (Nkrumah et al., 2007; Lu et al., 2013; Mao et al., 2013; Zhang et al., 2020). The RFI in kg of DMI per day was

computed as the difference between the standardized daily DMI and the expected DMI that was predicted based on animal ADG and MWT (Koch et al., 1963). Blood samples were collected from all animals by jugular venipuncture in the early morning on the first day of feedlot tests and immediately frozen at -80°C for storage. These blood samples were used to quantify metabolites using nuclear magnetic resonance (NMR) spectroscopy. The procedure of metabolite quantification using NMR was previously described by Li et al. (2020). Thirty-one metabolites and their concentration levels (μM) were quantified from plasma. Blood samples were also used to extract DNA for genotyping using the Illumina BovineSNP50 v2 BeadChip (Illumina Inc., CA, USA).

3.3.2 SNP genotype imputation, quality control and population admixture analyses

Theoretically, a higher marker density could improve the power of GWAS to identify significant SNPs, therefore, the 50K genotypes were imputed to whole genome sequence variants using Beagle 5.1 software (Browning et al., 2018). The SNP imputation for animals used in this study was completed using a step-wise approach as described by Zhang et al. (2020) and Wang et al. (2020) based on the latest genome assembly ARS-UCD 1.2. The reference populations were 4,247 animals of mixed beef breeds from our own reference population, and 3,093 *Bos taurus* animals from the 1000 Bull Genomes Project (Daetwyler et al., 2014) (run 7), respectively. After the imputation, 53,258,178 SNPs and indels (they are all termed SNPs for simplicity) on 29 autosomes were obtained. Quality control for imputed whole genome sequence variants was performed to exclude DNA variants based on the following criteria: SNPs on 29 autosomes that had an imputation accuracy < 0.95, minor allele frequency < 0.05, and failed to pass the Hardy-Weinberg equilibrium test (*P*-value < 0.0001). Finally, a total of 10,488,742 SNPs remained after quality control and were used in further analyses.

Breed composition of each animal was predicted based on the 50K genotypes using ADMIXTURE software to account for population stratification (Alexander et al., 2009; Hellwege et al., 2017). In order to find the best possible number of ancestors or breeds (K value), a 5-fold cross-validation procedure was performed as described in Zhang et al. (2020). The breed composition prediction had the smallest cross-validation error when the value of K = 6.

3.3.3 Data consolidation, quality control for feed efficiency traits and metabolites

The variation in feed efficiency traits and metabolites could be affected by multiple systematic effects. A linear regression model implemented in R statistical software was used to assess the significant systematic effects that were associated with feed efficiency traits or metabolites. Animal type (bull, heifer, steer), birth year, herd, feedlot pen, age at the feedlot test, and breeding composition were found to be the significantly associated factors for both the feed efficiency traits and metabolites (P-value < 0.05). Therefore, phenotypic values of the feed efficiency traits and metabolites were pre-adjusted for the above factors using liner regression models. Residuals with more or less than 3 standard deviations from the mean of residuals were considered as outliers and were excluded.

3.3.4 Regression analyses between feed efficiency traits and metabolites and metabolome-genome wide association studies

After quality control and pre-adjustment of phenotypic data, regression analyses were conducted to identify associations between four feed efficiency traits and thirty-one metabolites using R statistical software. A feed efficiency trait and a metabolite were considered to be significantly associated when a P-value < 0.1 of the regression analyses was observed. This step was intended to determine the relationship between feed efficiency traits and metabolites. The mGWAS (metabolome-genome wide association studies) were performed for metabolites that

were significantly associated with the feed efficiency traits using the mlma (mixed linear model association) option as implemented in the GCTA package (Yang et al., 2011) based on the following linear mixed model:

$$y_{ij} = \mu + b_j x_{ij} + a_{ij} + e_{ij}$$

where y_{ij} is the adjusted metabolite value of the ith animal with the jth SNP (i.e. the ijth animal), b_j is the allele substitution effect of the jth SNP, x_{ij} is the jth SNP genotype of animal i coded as 0, 1, 2 for genotypes A_1A_1, A_1A_2, and A_2A_2, respectively, a_{ij} is the additive polygenic effect of the ijth animal $\sim N(0, \boldsymbol{G}\sigma_a^2)$, and e_{ij} is the random residual effect $\sim N(0, \boldsymbol{I}\sigma_e^2)$. The genomic relationship matrix $\boldsymbol{G}$ that was derived based on total filtered SNP markers (i.e. 10,488,742 SNPs) as described by Yang et al. (2014), which is essentially the same as the second VanRaden formulation (VanRaden, 2008). The same $\boldsymbol{G}$ matrix was used to estimate variance components and heritability of metabolites via restricted maximum likelihood (REML) as implemented in the GCTA package.

The SNPs with *P*-value < 1×10^{-5} were classified to be significantly associated with the metabolite according to the recommendation of The Wellcome Trust Case Control Consortium (Burton et al., 2007). The phenotypic variance of the metabolite explained by each significant SNP was calculated by $\frac{2pq\beta^2}{S^2} * 100\%$, where p and q denote the SNP allele frequency of A_1 and A_2, respectively; β is the SNP allele substitute effect that was estimated by generalized least square and the significance of SNP allele substitution effect was conducted via a generalized least square F-test as implemented in the GCTA package; $2pq\beta^2$ is the additive variance of the SNP, and S^2 is the phenotypic variance of the metabolite.

3.3.5 Identification of candidate genes and functional enrichment analyses for feed efficiency traits

To identify candidate genes for concentration of each metabolite, a 140-kbp window (70-kbp upstream and 70-kbp downstream) of each significant SNP was surveyed based on SNP annotation information from ARS-UCD 1.2 bovine genome assembly from the Ensembl BioMart database (accessed on 02 February, 2021). The 70-kbp was the chromosomal length within which a high linkage disequilibrium phase correlation ($r^2 > 0.77$) was maintained across a sample of Canadian beef cattle breeds (Lu et al., 2012). Small nucleolar RNA and microRNA were excluded because we are interested in protein coding genes. Then candidate genes (Entrez gene IDs) of all metabolites that were associated with the feed efficiency traits (RFI, DMI, ADG, or MWT) as identified in the regression analyses were combined and imported into the Ingenuity Pathway Analysis software (accessed on 02 February, 2021) (IPA; www.Ingenuity.com) to predict the enriched biological functions and gene networks for feed efficiency traits. Biological functions were considered significantly enriched if the *P*-value for the overlap comparison test between the input gene list and the knowledge base of IPA for a given biological function was less than 0.05. In order to provide insight into cellular and molecular functions associated with feed efficiency traits, gene networks for some significant biological functions were constructed in IPA.

3.4 Results

3.4.1 Associations between feed efficiency traits and metabolites

Of the 31 metabolites analyzed, 11 were found to be significantly associated with the feed efficiency traits (*P*-value < 0.1) and the results of regression analyses are shown in Table 3.1. Among significant associated metabolites with each trait, ten metabolites showed *P*-values less than 0.05, and four metabolites (choline for DMI, glycine, citric acid and dimethyl sulfone for

MWT) showed *P*-values ranging from 0.05 to 0.1 (0.09, 0.05, 0.06, and 0.09, respectively). At *P*-values less than 0.1, five metabolites, including L-valine, lysine, L-tyrosine, L-isoleucine and L-leucine, were significantly associated with RFI, accounting for 5.90% of the phenotypic variance in RFI. Lysine, L-lactic acid, L-tyrosine and choline were significantly associated with DMI, and these four metabolites accounted for 4.04% of phenotypic variance in DMI. Of note, lysine and L-tyrosine were significantly associated with both RFI and DMI. Citric acid was the only metabolite that was significantly associated with ADG and accounted for 0.93% of phenotypic variance in ADG. Four metabolites, L-glutamine, glycine, citric acid and dimethyl sulfone, were significantly associated with MWT and accounted for 3.39% of phenotypic variance of MWT.

3.4.2 Significant SNPs and candidate genes associated with metabolites

Metabolome-genome wide association studies were performed for the 11 metabolites associated with the feed efficiency traits. The range of *P*-value and allele substitution effect of significant SNPs, the range and average of proportion of metabolite phenotypic variance explained by each significant SNP, and the number of quantitative trait loci (QTLs) and candidate genes identified for each metabolite are summarized in Table 3.2. In summary, 40, 66, 15 and 40 unique candidate genes were identified as related to RFI, DMI, ADG, and MWT, respectively (Table 3.3). Besides, 24 candidate genes were overlapped for RFI and DMI, 15 candidate genes were overlapped for ADG and MWT and 1 gene was common between DMI and MWT (Table 3.S1 and Figure 3.S1).

3.4.3 Significantly enriched biological functions and gene networks for feed efficiency traits

Of the 40, 66, 15 and 40 unique candidate genes, 39, 65, 15 and 39 genes for RFI, DMI, ADG and MWT were mapped to the IPA database for functional enrichment analyses, respectively. In summary, 24, 25, 18 and 28 significant cellular and molecular functions were identified for RFI,

DMI, ADG, and MWT (*P*-value < 0.05), respectively as presented in Table 3.S2 - 3.S5. The top five enriched cellular and molecular functions with corresponding candidate genes for each feed efficiency trait are shown in Table 3.4. Of the top five enriched cellular and molecular functions, lipid metabolism was the biological function with the lowest *P*-value for DMI and also significantly associated with RFI and MWT (Table 3.S3 and Table 3.S6). Molecular transport was one of the top five biological functions associated with DMI, ADG, and MWT. Small molecule biochemistry and nucleic acid metabolism were two top biological functions associated with both DMI and MWT. Among all significant biological functions, 15 biological functions were common for all four feed efficiency traits, and other biological functions shared among different feed efficiency traits are shown in Table 3.S6 and Figure 3.S2.

Additionally, in order to gain insight into potentially important biological functions, gene networks of lipid metabolism and carbohydrate metabolism were investigated and constructed through IPA. Within the lipid metabolism function for DMI, 16 candidate genes (*ACACB*, *ADGRF5*, *ALDH3B1*, *AQP9*, *CCDC80*, *CHKA*, *CPT1A*, *DAB1*, *DDX5*, *HNF1A*, *IGHMBP2*, *LRP5*, *NOS1*, *PLSCR1*, *PVALB* and *SSPN*) were involved (Figure 3.1). The lipid metabolism included nine subfunctions which were concentration of fatty acid, concentration of lipid, concentration of phosphatidylcholine, concentration of triacylglycerol, fatty acid metabolism, cholesterol metabolism, synthesis of fatty acid, synthesis of lipid, and transport of lipid (Figure 3.1). Interestingly, seven genes (*ACACB*, *AQP9*, *CCDC80*, *CHKA*, *CPT1A*, *HNF1A* and *LRP5*) involved in the lipid metabolism were also involved in the carbohydrate metabolism for DMI, which engaged three subfunctions including oxidation of D-glucose, concentration of D-glucose and quantity of carbohydrate (Figure 3.2).

3.5 Discussion

3.5.1 The role of metabolites in variation of feed efficiency traits

Variation in RFI and its component traits could represent differences among animals in terms of metabolic process activity. For example, a study has shown that low RFI steers tend to have more efficient metabolic process activity and are able to meet their maintenance requirement with less energy intake than high RFI steers (Richardson and Herd, 2004). Blood is the major highway for transportation of nutrients to the different organs and tissues, and metabolites carried by blood are directly involved in metabolic processes as substrates or products, making blood metabolites prime candidates for further studies of feed efficiency in beef cattle. Additionally, some blood metabolites have the potential to serve as biomarkers for selection of efficient beef cattle (Karisa et al., 2014; Novais et al., 2019).

In this study, 5 (L-valine, lysine, L-tyrosine, L-isoleucine, and L-leucine), 4 (lysine, L-lactic acid, L-tyrosine, and choline), 1 (citric acid) and 4 (L-glutamine, glycine, citric acid, and dimethyl sulfone) plasma metabolites were identified to be associated with RFI, DMI, ADG, and MWT, respectively (Table 3.1). Individual metabolites accounted for 0.59% to 1.50% of the total phenotypic variance of RFI and its component traits. The results suggest that the feed efficiency traits could be associated with many metabolites with small effects. However, the identified metabolites associated with the feed efficiency traits in this study may require validation in independent beef cattle populations especially as a more relaxed threshold (P-value < 0.1) was used. Furthermore, we would like to highlight that only 31 metabolites were detected by the targeted method of NMR used in the current study. We therefore recommend that metabolomic profiles with more metabolites should be investigated in future with larger numbers of samples to identify more metabolites that are associated with RFI or its component traits.

To date, several metabolomic studies have attempted to identify relationships between serum or plasma metabolite levels and RFI in beef cattle (Karisa et al., 2014; Jorge-Smeding et al., 2019; Foroutan et al., 2020). We found good agreement between the results from those studies and the current study. In the current study, valine, lysine, tyrosine, and leucine showed higher concentrations in beef cattle with high RFI than those with low RFI. In line with our results, Karisa et al. (2014) and Foroutan et al. (2020) observed higher concentrations of valine, lysine, and tyrosine in beef cattle with high RFI as compared to those with low RFI. Similarly, Jorge-Smeding et al. (2019) reported that concentrations of valine and lysine were decreased in heifers with low RFI. Additionally, Foroutan et al. (2020) reported the concentration of leucine was higher in high-RFI beef cattle, which is consistent with our results. The consistency of results from different studies suggests that these metabolites have the potential to be used as biomarkers for feed efficiency.

It is worth noting that, the three metabolites (isoleucine, leucine, and valine) associated with RFI are three essential branched-chain amino acids. These three metabolites share the first enzymatic steps in their oxidative pathways, including a reversible transamination followed by an irreversible oxidative decarboxylation to coenzyme-A derivatives (Manoli and Venditti, 2016). The respective oxidative pathways subsequently diverge and at the final steps yield acetyl- and/or propionyl-CoA that enter the tricarboxylic acid cycle (TCA cycle) (Manoli and Venditti, 2016). For animals, the TCA cycle is the main energy producing (mainly from carbohydrates and fatty acids) metabolic pathway (Akram, 2014), and some of the processes of the TCA cycle pathway have been previously reported to be associated with feed efficiency in beef cattle (Karisa et al., 2014) and pigs (Wang and Kadarmideen, 2019). Additionally, in this study, citric acid was the only metabolite that was significantly associated with ADG and was overlapping for ADG and

MWT. Citric acid is an important intermediate in the TCA cycle (Akram, 2014) indicating a potential relationship between the TCA cycle related metabolic processes and feed efficiency traits. Interestingly, two other metabolites (lysine and L-tyrosine) were identified as associated with both RFI and DMI in this study. In the current study, we observed that the concentrations of lysine and L-tyrosine were significantly positively correlated (r = 0.29, P-value < 0.001). A previous study reported a higher positive correlation (r > 0.75, P-value < 0.001) between lysine and tyrosine (Wang and Kadarmideen, 2019). The association of lysine and L-tyrosine with both RFI and DMI could be due to the significant positive correlation between lysine and tyrosine and the fact that RFI has a high and positive genetic correlation with DMI (r_g = 0.66 ± 0.11 to 0.75 ± 0.10) (Mao et al., 2013; Ceacero et al., 2016). Furthermore, lysine and tyrosine were reported as important amino acids involved in some important metabolic processes in beef cattle, such as amino acid metabolism and urea cycle (Jorge-Smeding et al., 2019), further supporting them as potential biomarkers for feed efficiency traits.

3.5.2 Candidate genes, enriched molecular functions and gene networks for feed efficiency traits

In this study, we identified 40, 66, 15 and 40 unique candidate genes as related to RFI, DMI, ADG, and MWT respectively via integrative analyses of regression analyses and mGWAS (Table 3.3). In a previous study, Zhang et al. (2020) performed GWAS based on imputed whole genome sequence variants for RFI, DMI, ADG, and MWT using 7,500 beef cattle and reported 596, 268, 179 and 532 candidate genes for RFI, DMI, ADG, and MWT, respectively. Comparing their results with those in this study, we found 10, 23, 6 and 7 candidate genes in common between the two studies for RFI, DMI, ADG, and MWT, respectively (Table 3.S7). These overlapping genes indicated that metabolites are potentially important intermediate phenotypes between

candidate genes and feed efficiency traits. Additionally, results from our study provide more knowledge and better understanding of how the previously identified candidate genes exert their influence on the variability of RFI and its component traits via intermediate phenotype metabolites. For instance, Zhang et al. (2020) reported that some genes were associated with more than one trait such as, *ADGRF1* and *ADGRF5* which were associated with both RFI and DMI (Zhang et al., 2020). However, the potential mechanism of how these genes could influence the two traits remained unclear. According to the results of the current study, these two genes were both associated with L-tyrosine as a common metabolite which was associated with RFI and DMI (Table 3.3). Similarly, according to Zhang et al. (2020), *SLC28A3* was associated with ADG and MWT, and our results showed this gene was associated with citric acid as a common metabolite which was associated with ADG and MWT (Table 3.3). Interestingly, Zhang et al. (2020) identified *ADGRF1*, *ADGRF5*, *GTPBP8* and *NEPRO* as associated with both RFI and DMI and the same genes for RFI and DMI were identified in the current study. However, the results of this study indicated that the molecular background of these associations might be different. L-tyrosine might explain the associations of *ADGRF1*, *ADGRF5* with RFI and DMI, because we identified that *ADGRF1* and *ADGRF5* were associated with L-tyrosine which was a metabolite associated with both RFI and DMI. As for *GTPBP8* and *NEPRO*, both genes were associated with another common metabolite, lysine that was identified to be associated with both RFI and DMI in the current study. Additionally, we observed that certain genes might be associated with the same feed efficiency trait through different metabolites. For example, *SHROOM3* was associated with L-valine and L-leucine and these two metabolites were associated with RFI (Table 3.3). Our study also showed that certain genes could be associated with different feed efficiency traits through different metabolites. For example, *AQP9* was associated with DMI and MWT through L-lactic

acid and glycine, respectively (Table 3.3). Therefore, our integrative analyses of feed efficiency traits, metabolites, and whole genome sequence variants will enhance our understanding on genetic influence of feed efficiency traits in beef cattle.

Some candidate genes identified for feed efficiency traits in the current study have been reported in our previous transcriptomic studies involving animals related to those used in the current study (Mukiibi et al., 2018, 2019). For instance, *CCDC80* was reported as a differentially expressed gene between beef steers with divergent RFI (Mukiibi et al., 2018). Additionally, *CCDC80*, *CUX2* and *ALDH3B1* were differentially expressed in the liver of beef steers for DMI, and *SERPINE3* was a differentially expressed gene for ADG (Mukiibi et al., 2019). Our current study identified the same genes associated with these traits through integrating metabolites (Table 3.3). Indeed, *CCDC80*, *CUX2*, *ALDH3B1* and *SERPINE3* were associated with lysine, L-lactic acid, choline and citric acid, respectively. Therefore, our results potentially provide further insight into how these differentially expressed genes affect the feed efficiency traits in beef cattle. It is worth noting that *CUX2* has also been reported to be associated with DMI in the American (Seabury et al., 2017) and Canadian beef population (Zhang et al., 2020). Therefore, these genes identified as associated with the same feed efficiency traits using genomic, transcriptomic and metabolomic data suggest the importance of these genes in influencing feed efficiency traits in beef cattle. Furthermore, some differentially expressed genes may affect RFI by influencing metabolites associated with its component traits (DMI, ADG, and MWT). For example, *TCIRG1*, *AMN,* and *AQP9* were reported as differentially expressed genes in high- and low-RFI beef cattle (Tizioto et al., 2015; Weber et al., 2016) and these three genes were identified to be respectively associated with DMI, ADG, and MWT through different metabolites in this study.

Identification of enriched molecular processes, pathways and gene networks associated with feed efficiency traits using candidate genes from these different omics studies shed some light on underlying biological mechanism and gene interactions for complex traits. For the five topmost biological functions associated with RFI in the current study, cellular assembly and organization, cell morphology, cellular function and maintenance, and molecular transport were four biological functions that overlapped with the five topmost biological functions reported by Zhang et al. (2020) for RFI. Lipid metabolism, small molecule biochemistry and nucleic acid metabolism were three common top biological functions for DMI in the two studies. Lipid metabolism and small molecule biochemistry were also identified as two of the five topmost biological functions in our previous transcriptomic study for DMI in beef cattle (Mukiibi et al., 2019). Molecular transport, small molecule biochemistry and cell morphology were three overlapping top biological functions for MWT in Zhang et al. (2020) and this study. These three biological functions were also top biological functions for MWT in our previous transcriptomic study (Mukiibi et al., 2019). For ADG, cell-to-cell signaling and interaction was a common top biological functions in Zhang et al. (2020), Mukiibi et al. (2019) and the current study. Our results and those reported by previous studies indicated the overlapping top five biological functions have a potentially important relationship with feed efficiency traits in beef cattle. These important functions could further help to prioritize candidate genes and related functional SNPs associated with phenotypes.

Additionally, we would like to note that attention should be paid to nutrient or energy metabolic processes, such as lipid metabolism, since several studies have reported its potential role in feed efficiency related to DMI and RFI (Nkrumah et al., 2007; Chen et al., 2011; Mao et al., 2013; Alexandre et al., 2015; Weber et al., 2016; McKenna et al., 2018; Mukiibi et al., 2018, 2019; Higgins et al., 2019; Zhang et al., 2020). Nkrumah et al. (2007) and Mao et al. (2013) reported that

more efficient beef cattle tended to have less backfat and slightly less marbling. Transcriptomic studies reported that more efficient beef cattle were associated with differentially expressed genes related to reducing lipid metabolism in liver (Mukiibi et al., 2018, 2019), implying an important relationship between lipid metabolism and feed efficiency. Weber and colleagues identified differentially expressed genes in multiple tissues (pituitary, skeletal muscle, liver, visceral adipose, and duodenum) of beef cattle with divergent RFI, and their pathway analyses showed that many of the differentially expressed genes were involved in the immune system and fat metabolism (Weber et al., 2016). In this study, lipid metabolism was the most significant biological function for DMI and it was also significantly associated with RFI and MWT. Lipid metabolism was identified as one of the top biological functions for ADG in previous studies (Mukiibi et al., 2019; Zhang et al., 2020) but it was not shown in the current study, which is likely due to limitations of relatively small number of metabolites analyzed. In addition, of the candidate genes identified for the metabolites, there is limited knowledge on how candidate genes influence the respective plasma metabolite levels. For instance, enzyme choline kinase alpha is encoded by *CHKA* (Hosaka et al., 1992). In the biosynthesis pathway of phosphatidylcholine, the enzyme can catalyze the phosphorylation of choline to phosphocholine (Lacal, 2001; Aoyama et al., 2004). However, little is known on how concentrations of choline vary among animals due to their gene variants. Nevertheless, our integrative study of feed efficiency, blood metabolites and DNA variants has provided additional insight into relationships between gene functionalities, metabolites and feed efficiency traits, which may help develop strategies to enhance genomic prediction of feed efficiency traits with incorporation of metabolite data.

3.6 Conclusion

This study combined genomic, metabolomic and phenotypic data to investigate molecules and biological functions/processes related to feed efficiency in beef cattle. Several plasma metabolites associated with RFI and its component traits were identified, and some of metabolites showed the potential to serve as biomarkers for feed efficiency in beef cattle. Multiple candidate genes were identified as associated with RFI and its component traits based on the results of regression analyses between feed efficiency traits and metabolites, and mGWAS. Gene functional enrichment analyses indicated that lipid metabolism may have an important role in feed efficiency. Our findings showed good consistency with previous metabolomic studies and GWAS studies for feed efficiency and also added more information regarding biological mechanisms of feed efficiency. Therefore, this integrative method could enhance the understanding of genetic influence, metabolites and biological functions/processes involved in feed efficiency traits, which could lead to improvement of genomic prediction accuracy via incorporating metabolite data.

3.7 References

Abo-Ismail, M. K., Lansink, N., Akanno, E., Karisa, B. K., Crowley, J. J., Moore, S. S., et al. (2018). Development and validation of a small SNP panel for feed efficiency in beef cattle. *J. Anim. Sci.* 96, 375–397. doi:10.1093/jas/sky020.

Abo-Ismail, M. K., Vander Voort, G., Squires, J. J., Swanson, K. C., Mandell, I. B., Liao, X., et al. (2014). Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. *BMC Genet.* 15, 14. doi:10.1186/1471-2156-15-14.

Ahola, J. K., and Hill, R. A. (2012). "Input Factors Affecting Profitability: A Changing Paradigm and a Challenging Time," in *Feed Efficiency in the Beef Industry* (Hoboken, New Jersey: Wiley-Blackwell), 7–19. doi:10.1002/9781118392331.ch1.

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle - An independent validation study. *J. Anim. Sci.* 92, 2896–2904. doi:10.2527/jas.2013-7374.

Akram, M. (2014). Citric acid cycle and role of its intermediates in metabolism. *Cell Biochem. Biophys.* 68, 475–478. doi:10.1007/s12013-013-9750-1.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109.

Alexandre, P. A., Kogelman, L. J. A., Santana, M. H. A., Passarelli, D., Pulz, L. H., Fantinato-Neto, P., et al. (2015). Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16, 1073. doi:10.1186/s12864-015-2292-8.

Aoyama, C., Liao, H., and Ishidate, K. (2004). Structure and function of choline kinase isoforms in mammalian cells. *Prog. Lipid Res.* 43, 266–281. doi:10.1016/j.plipres.2003.12.001.

Archer, J. A., Barwick, S. A., and Graser, H. U. (2004). Economic evaluation of beef cattle breeding schemes incorporating performance testing of young bulls for feed intake. *Aust. J. Exp. Agric.* 44, 393–404. doi:10.1071/ea02054.

Arthur, P. F., Renand, G., and Krauss, D. (2001). Genetic and phenotypic relationships among different measures of growth and feed efficiency in young Charolais bulls. *Livest. Prod. Sci.* 68, 131–139. doi:10.1016/S0301-6226(00)00243-8.

Bolormaa, S., Hayes, B. J., Savin, K., Hawken, R., Barendse, W., Arthur, P. F., et al. (2011). Genome-wide association studies for feedlot and growth traits in cattle. *J. Anim. Sci.* 89, 1684–1697. doi:10.2527/jas.2010-3079.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi:10.1016/j.ajhg.2018.07.015.

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678. doi:10.1038/nature05911.

Ceacero, T. M., Mercadante, M. E. Z., Cyrillo, J. N. D. S. G., Canesin, R. C., Bonilha, S. F. M.,

and De Albuquerque, L. G. (2016). Phenotypic and genetic correlations of feed efficiency traits with growth and carcass traits in nellore cattle selected for postweaning weight. *PLoS One* 11, e0161366. doi:10.1371/journal.pone.0161366.

Chen, Y., Gondro, C., Quinn, K., Herd, R. M., Parnell, P. F., and Vanselow, B. (2011). Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim. Genet.* 42, 475–490. doi:10.1111/j.1365-2052.2011.02182.x.

Fontanesi, L. (2016). Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. *Anim. Front.* 6, 73. doi:10.2527/af.2016-0011.

Foroutan, A., Fitzsimmons, C., Mandal, R., Berjanskii, M. V., and Wishart, D. S. (2020). Serum metabolite biomarkers for predicting residual feed intake (RFI) of young angus bulls. *Metabolites* 10, 491. doi:10.3390/metabo10120491.

Gerber, P. J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., et al. (2013). *Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities*. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).

Hegarty, R. S., Goopy, J. P., Herd, R. M., and McCorkell, B. (2007). Cattle selected for lower residual feed intake have reduced daily methane production. *J. Anim. Sci.* 85, 1479–1486. doi:10.2527/jas.2006-236.

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., and Edwards, T. L. (2017). Population stratification in genetic association studies. *Curr. Protoc. Hum. Genet.* 95, 1.22.1-1.22.23. doi:10.1002/cphg.48.

Herd, R. M., and Bishop, S. C. (2000). Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. *Livest. Prod. Sci.* 63, 111–119. doi:10.1016/S0301-6226(99)00122-0.

Higgins, M. G., Kenny, D. A., Fitzsimons, C., Blackshields, G., Coyle, S., McKenna, C., et al. (2019). The effect of breed and diet type on the global transcriptome of hepatic tissue in beef cattle divergent for feed efficiency. *BMC Genomics* 20, 525. doi:10.1186/s12864-019-5906-8.

Hosaka, K., Tanaka, S., Nikawa, J. ichi, and Yamashita, S. (1992). Cloning of a human choline kinase cDNA by complementation of the yeast *cki* mutation. *FEBS Lett.* 304, 229–232. doi:10.1016/0014-5793(92)80625-Q.

Jorge-Smeding, E., Renand, G., Centeno, D., Pétéra, M., Durand, S., Polakof, S., et al. (2019). Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. in *EAAP Scientific Series* (Wageningen, Netherlands: Wageningen Academic Publishers), 231–232. doi:10.3920/978-90-8686-891-9_50.

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., et al. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livest. Sci.* 165, 200–211. doi:10.1016/j.livsci.2014.03.002.

Koch, R. M., Swiger, L. A., Chambers, D., and Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *J. Anim. Sci.* 22, 486–494. doi:10.2527/jas1963.222486x.

Lacal, J. C. (2001). Choline kinase: a novel target for antitumor drugs. *IDrugs* 4, 419–26.

doi:10.1042/bj1800559.

Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., et al. (2020). Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle. *Front. Genet.* 11, 538600. doi:10.3389/fgene.2020.538600.

Lu, D., Miller, S., Sargolzaei, M., Kelly, M., Vander Voort, G., Caldwell, T., et al. (2013). Genome-wide association analyses for growth and feed efficiency traits in beef cattle. *J. Anim. Sci.* 91, 3612–3633. doi:10.2527/jas.2012-5716.

Lu, D., Sargolzaei, M., Kelly, M., Li, C., Voort, G. Vander, Wang, Z., et al. (2012). Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Genet.* 3, 152. doi:10.3389/fgene.2012.00152.

Manoli, I., and Venditti, C. P. (2016). Disorders of branched chain amino acid metabolism. *Transl. Sci. Rare Dis.* 1, 91–110. doi:10.3233/trd-160009.

Mao, F., Chen, L., Vinsky, M., Okine, E., Wang, Z., Basarab, J., et al. (2013). Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci.* 91, 2067–2076. doi:10.2527/jas.2012-5470.

McKenna, C., Porter, R. K., Keogh, K. A., Waters, S. M., McGee, M., and Kenny, D. A. (2018). Residual feed intake phenotype and gender affect the expression of key genes of the lipogenesis pathway in subcutaneous adipose tissue of beef cattle. *J. Anim. Sci. Biotechnol.* 9, 68. doi:10.1186/s40104-018-0282-9.

McKeown, L., Aalhus, J., Larsen, I., Stothard, P., and Wang, Z. (2013). Bridging the "Phenomic Gap": Creation of a database containing phenotypes and genotypes for economically important traits for beef cattle. Edmonton, Canada. Final Report to the Alberta Livestock and Meat Agency.

Mukiibi, R., Vinsky, M., Keogh, K. A., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2018). Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. *Sci. Rep.* 8, 7303. doi:10.1038/s41598-018-25605-3.

Mukiibi, R., Vinsky, M., Keogh, K., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2019). Liver transcriptome profiling of beef steers with divergent growth rate, feed intake, or metabolic body weight phenotypes. *J. Anim. Sci.* 97, 4386–4404. doi:10.1093/jas/skz315.

Nielsen, M. K., MacNeil, M. D., Dekkers, J. C. M., Crews, D. H., Rathje, T. A., Enns, R. M., et al. (2013). Life-cycle, total-industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation11The development of this commentary was supported by the Beef Improvement Federation. *Prof. Anim. Sci.* 29, 559–565. doi:10.15232/S1080-7446(15)30285-0.

Nkrumah, J. D., Basarab, J. A., Wang, Z., Li, C., Price, M. A., Okine, E. K., et al. (2007). Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. *J. Anim. Sci.* 85, 2711–2720. doi:10.2527/jas.2006-767.

Nkrumah, J. D., Okine, E. K., Mathison, G. W., Schmid, K., Li, C., Basarab, J. A., et al. (2006). Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J. Anim. Sci.* 84, 145–153.

doi:10.2527/2006.841145x.

Novais, F. J., Pires, P. R. L., Alexandre, P. A., Dromms, R. A., Iglesias, A. H., Ferraz, J. B. S., et al. (2019). Identification of a metabolomic signature associated with feed efficiency in beef cattle. *BMC Genomics* 20, 8. doi:10.1186/s12864-018-5406-2.

Ramsey, R., Doye, D., Ward, C., McGrann, J., Falconer, L., and Bevers, S. (2005). Factors affecting beef cow-herd costs, production, and profits. *J. Agric. Appl. Econ.* 37, 91–99. doi:10.1017/s1074070800007124.

Richardson, E. C., and Herd, R. M. (2004). Biological basis for variation in residual feed intake in beef cattle. 2. Synthesis of results following divergent selection. *Aust. J. Exp. Agric.* 44, 431–440. doi:10.1071/ea02221.

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., et al. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18, 386. doi:10.1186/s12864-017-3754-y.

Tizioto, P. C., Coutinho, L. L., Decker, J. E., Schnabel, R. D., Rosa, K. O., Oliveira, P. S. N., et al. (2015). Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. *BMC Genomics* 16, 242. doi:10.1186/s12864-015-1464-x.

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

Wang, X., and Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites* 9, 151. doi:10.3390/metabo9070151.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Weber, K. L., Welly, B. T., Van Eenennaam, A. L., Young, A. E., Porto-Neto, L. R., Reverter, A., et al. (2016). Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. *PLoS One* 11, e0152274. doi:10.1371/journal.pone.0152274.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi:10.1016/j.ajhg.2010.11.011.

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46, 100–106. doi:10.1038/ng.2876.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (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, 36. doi:10.1186/s12864-019-6362-1.

Table 3.1 A summary of metabolites associated with RFI and its component traits in a multibreed population of beef cattle

Trait[1]	Metabolite[2]	*P*-value[3]	b[4]	V_m/V_P (%)[5]
RFI	L-valine	6.94E-03	2.72E-03	5.90
	Lysine	9.61E-03	3.96E-03	
	L-tyrosine	2.40E-02	6.65E-03	
	L-isoleucine	2.64E-02	5.80E-03	
	L-leucine	3.40E-02	3.13E-03	
DMI	Lysine	1.15E-02	5.06E-03	4.04
	L-lactic acid	2.25E-02	-6.98E-05	
	L-tyrosine	2.45E-02	8.69E-03	
	choline	9.27E-02	6.69E-04	
ADG	citric acid	3.56E-02	4.31E-04	0.93
MWT	L-glutamine	1.49E-02	3.57E-02	3.39
	glycine	5.29E-02	-4.79E-03	
	citric acid	6.11E-02	-9.56E-03	
	dimethyl sulfone	9.67E-02	-2.63E-02	

[1]*RFI* residual feed intake in kg of DMI per day, *DMI* daily dry matter intake in kg per day, *ADG* average daily gain in kg, *MWT* metabolic body weight in kg

[2]The unit of metabolite concentration is μM

[3]The significance level of regression analysis is *P*-value < 0.1

[4]*b* regression coefficient

[5]V_m/V_P the proportion of phenotypic variance of feed efficiency traits explained by associated metabolites (%)

Table 3.2 A summary of significant SNPs, the number of QTLs, and the number of candidate genes for metabolites associated with feed efficiency traits in a multibreed population of beef cattle

Metabolite[1]	*P*-value range[2]	β range[3]	V_{SNP}/V_P range (%)[4]	V_{SNP}/V_P mean (%)[5]	No. of QTL[6]	No. of gene[7]
citric acid	1.47E-06 – 9.75E-06	-29.80 – 37.62	3.57 – 4.95	4.05	15	15
choline	4.94E-07 – 9.90E-06	-89.13 – 84.25	3.85 – 5.43	4.61	13	23
glycine	3.17E-06 – 9.54E-06	68.80 – 75.32	3.97 – 4.65	4.31	9	10
L-tyrosine	2.75E-06 – 9.40E-06	-4.20 – 7.73	3.96 – 4.49	4.11	5	2
L-isoleucine	4.21E-06 – 8.94E-06	-8.88 – 9.75	4.04 – 4.35	4.14	3	3
lysine	9.11E-09 – 9.80E-06	-17.77 – 20.56	3.88 – 7.13	4.82	15	20
L-lactic acid	2.24E-07 – 9.43E-06	-1076.62 – 1261.24	3.74 – 5.95	4.58	16	21
L-glutamine	7.37E-07 – 9.90E-06	-11.76 – 11.53	4.06 – 5.30	4.66	13	13
L-leucine	1.03E-06 – 9.30E-06	-18.12 – 17.33	3.99 – 5.04	4.40	9	12
L-valine	3.64E-06 – 9.78E-06	-25.43 – 24.38	3.73 – 4.46	4.03	8	4
dimethyl sulfone	9.44E-07 – 9.53E-06	-8.46 – 7.52	3.90 – 5.04	4.45	4	2

[1]The unit of metabolite concentration is μM

[2]The *P*-value range (minimum to maximum) of significant SNPs, the significance level is *P*-value < 1×10^{-5}

[3]*β range* the range of allele substitution effect of each significant SNP

[4]*V_{SNP}/V_P range* the range of metabolite phenotypic variance explained by each significant SNP (%)

[5]*V_{SNP}/V_P mean* the average of metabolite phenotypic variance explained by each significant SNP (%)

[6]*No. of QTL* the number of QTLs identified for each metabolite

[7]*No. of gene* the number of candidate gene identified for each metabolite

Table 3.3 Metabolites and their candidate genes associated with RFI and its component traits in a multibreed population of beef cattle

Trait[1]	Metabolite[2]	Candidate gene
RFI	L-valine	*NEDD4, PRTG, SHROOM3, XKR6*
	lysine	*BTLA, ATG3, SLC35A5, CCDC80, CD200R1L, GTPBP8, NEPRO, BOC, SPICE1, SIDT1, FGF12, HS6ST3, FRMD5, MFHAS1, STYXL2, GPA33, DAB1, OR6C75, ITPR2, SSPN*
	L-tyrosine	*ADGRF5, ADGRF1*
	L-isoleucine	*C15H11orf49, PPYR1, ANXA8L1*
	L-leucine	*SLC9A9, SYNE2, ESR2, DYNC1LI1, CD2AP, ADGRF2, ADGRF4, SHROOM3, KATNA1, LATS1, NUP43, PCMT1*
DMI	lysine	*BTLA, ATG3, SLC35A5, CCDC80, CD200R1L, GTPBP8, NEPRO, BOC, SPICE1, SIDT1, FGF12, HS6ST3, FRMD5, MFHAS1, STYXL2, GPA33, DAB1, OR6C75, ITPR2, SSPN*
	L-lactic acid	*PLSCR1, AQP9, NEDD4, PRTG, PYGO1, CUX2, NOS1, FBXO21, SPPL3, HNF1A, C17H12orf43, OASL, FOXN4, ACACB, TMEM171, FCHO2, CD247, POU2F1, MACF1, NPFFR2, SGCD*
	L-tyrosine	*ADGRF5, ADGRF1*
	choline	*HHAT, CDH8, PECAM1, MILR1, POLG2, DDX5, CEP95, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, LRP5, PPP6R3, CPT1A, MRPL21, IGHMBP2, MRGPRF, CACNG2, IFT27, PVALB, BICD1, PERP*
ADG	citric acid	*SERPINE3, INTS6, ZNF667, ZNF583, USP32, CA4, ZNHIT3, MYO19, TRAF3, AMN, CDC42BPB, EDEM1, ARL8B, KLHL31, SLC28A3*
MWT	L-glutamine	*MYO16, UBE2E2, DDX56, NPC1L1, NUDCD3, CAMK2B, TRIM24, SVOPL, ATP6V0A4, PPP3CC, SORBS3, PDLIM2, CCAR2*
	glycine	*AQP9, PHLDB1, TREH, DDX6, EIF5, MARK3, SEM1, PINX1, SOX7, C8H8orf74*
	citric acid	*SERPINE3, INTS6, ZNF667, ZNF583, USP32, CA4, ZNHIT3, MYO19, TRAF3, AMN, CDC42BPB, EDEM1, ARL8B, KLHL31, SLC28A3*
	dimethyl sulfone	*ULK4, TRAK1*

[1]*RFI* residual feed intake in kg of DMI per day, *DMI* daily dry matter intake in kg per day, *ADG* average daily gain in kg, *MWT* metabolic body weight in kg

[2]The unit of metabolite concentration is μM

Table 3.4 Five topmost significantly enriched biological functions for RFI and its component traits, and genes involved in functions

Trait[1]	Biological function	*P*-value range[2]	Genes involved in the biological function
RFI	Cellular Assembly and Organization	7.92E-05 – 4.19E-02	*ADGRF1, ADGRF5, ANXA8L1, ATG3, BOC, CD2AP, DAB1, DYNC1LI1, ESR2, FGF12, ITPR2, KATNA1, LATS1, NEDD4, SHROOM3, SPICE1, SYNE2*
	Cell Morphology	1.02E-03 – 4.02E-02	*ADGRF5, ATG3, BOC, CD2AP, ESR2, KATNA1, LATS1, NEDD4, SLC9A9, SYNE2*
	Cellular Function and Maintenance	1.02E-03 – 4.19E-02	*ADGRF1, ADGRF5, ANXA8L1, ATG3, BOC, CD2AP, DAB1, DYNC1LI1, ESR2, FGF12, ITPR2, KATNA1, NEDD4, SHROOM3, SYNE2*
	Cellular Movement	1.12E-03 – 2.87E-02	*DAB1, ESR2, KATNA1*
	Molecular Transport	1.28E-03 – 3.36E-02	*ADGRF5, DAB1, ESR2, ITPR2, LATS1, SHROOM3*
DMI	Lipid Metabolism	2.46E-04 – 2.81E-02	*ACACB, ADGRF5, ALDH3B1, AQP9, CCDC80, CHKA, CPT1A, DAB1, DDX5, HNF1A, IGHMBP2, LRP5, NOS1, PLSCR1, PVALB, SSPN*
	Molecular Transport	2.46E-04 – 2.54E-02	*ACACB, ADGRF5, AQP9, CCDC80, CD247, CHKA, CPT1A, DAB1, DDX5, FGF12, HNF1A, IFT27, IGHMBP2, ITPR2, LRP5, NEDD4, NOS1, NPFFR2, PECAM1, PLSCR1, PVALB, TCIRG1*
	Small Molecule Biochemistry	2.46E-04 – 2.81E-02	*ACACB, ADGRF5, ALDH3B1, AQP9, CCDC80, CHKA, CPT1A, DAB1, DDX5, HNF1A, HS6ST3, IGHMBP2, LRP5, NOS1, NPFFR2, PECAM1, PLSCR1, PVALB, SGCD, SSPN, TCIRG1*
	Nucleic Acid Metabolism	2.84E-04 – 2.81E-02	*ACACB, CPT1A, NOS1*
	Protein Synthesis	5.66E-04 – 1.13E-02	*ACACB, CCDC80, HNF1A, LRP5, NOS1, NPFFR2, PECAM1, SGCD*
ADG	Cell-To-Cell Signaling and Interaction	6.57E-04 – 6.57E-04	*TRAF3*
	Cellular Development	6.57E-04 – 2.34E-02	*TRAF3*
	Cellular Function and Maintenance	6.57E-04 – 2.21E-02	*AMN, ARL8B, MYO19, TRAF3*
	Cellular Growth and Proliferation	6.57E-04 – 2.34E-02	*TRAF3*

	Molecular Transport	6.57E-04 – 2.34E-02	*AMN, ARL8B, CA4, SLC28A3, TRAF3*
MWT	Molecular Transport	7.04E-04 – 4.79E-02	*AMN, AQP9, ARL8B, ATP6V0A4, CA4, CAMK2B, DDX56, DDX6, NPC1L1, PPP3CC, SLC28A3, SORBS3, TRAF3, TRAK1*
	Nucleic Acid Metabolism	7.04E-04 – 2.7E-02	*AQP9, SLC28A3*
	Small Molecule Biochemistry	7.04E-04 – 4.79E-02	*AMN, AQP9, NPC1L1, PPP3CC, SLC28A3, SORBS3, TREH*
	Cell Cycle	1.45E-03 – 2.78E-02	*ARL8B, CAMK2B, MYO19, NUDCD3, TRIM24*
	Cell Morphology	1.45E-03 – 3.36E-02	*AQP9, ARL8B, CAMK2B, NUDCD3, PDLIM2, PINX1, TRIM24*

[1]*RFI* residual feed intake in kg of DMI per day, *DMI* daily dry matter intake in kg per day, *ADG* average daily gain in kg, *MWT* metabolic body weight in kg

[2]The *P*-value range (minimum to maximum) of significant biological functions, the significance level is *P*-value < 0.05

Table 3.S1 Uniquely common candidate genes for feed efficiency traits in a beef cattle multibreed population

Trait	Number of common genes	Common gene name
DMI, RFI	24	*ADGRF1, FRMD5, OR6C75, PRTG, SLC35A5, CCDC80, ATG3, GTPBP8, STYXL2, SIDT1, HS6ST3, BOC, NEPRO, BTLA, GPA33, DAB1, NEDD4, FGF12, MFHAS1, ITPR2, CD200R1L, SSPN, SPICE1, ADGRF5*
ADG, MWT	15	*ZNF667, CDC42BPB, CA4, SERPINE3, ZNHIT3, INTS6, KLHL31, SLC28A3, EDEM1, ARL8B, AMN, TRAF3, ZNF583, MYO19, USP32*
DMI, MWT	1	*AQP9*

Table 3.S2 Enriched biological functions significantly associated with RFI in a beef cattle multibreed population

Biological function	*P*-value range	Genes involved in the biological function
Cellular Assembly and Organization	7.92E-05-4.19E-02	*ADGRF1,ADGRF5,ANXA8/ANXA8L1,ATG3, BOC,CD2AP,DAB1,DYNC1LI1,ESR2,FGF12, ITPR2,KATNA1,LATS1,NEDD4,SHROOM3, SPICE1,SYNE2*
Cell Morphology	1.02E-03-4.02E-02	*ADGRF5,ATG3,BOC,CD2AP,ESR2,KATNA1, LATS1,NEDD4,SLC9A9,SYNE2*
Cellular Function and Maintenance	1.02E-03-4.19E-02	*ADGRF1,ADGRF5,ANXA8/ANXA8L1,ATG3, BOC,CD2AP,DAB1,DYNC1LI1,ESR2,FGF12, ITPR2,KATNA1,NEDD4,SHROOM3,SYNE2*
Cellular Movement	1.12E-03-2.87E-02	*DAB1,ESR2,KATNA1*
Molecular Transport	1.28E-03-3.36E-02	*ADGRF5,DAB1,ESR2,ITPR2,LATS1,SHROOM3*
Protein Synthesis	1.28E-03-3.2E-02	*ESR2,LATS1,SHROOM3*
Small Molecule Biochemistry	1.28E-03-4.31E-02	*ADGRF5,DAB1,ESR2,HS6ST3,LATS1,SSPN*
Cell-To-Cell Signaling and Interaction	1.71E-03-3.7E-02	*CD200R1L,ESR2,FGF12,LATS1,NEDD4,SLC9A9*
Cellular Development	1.71E-03-4.46E-02	*BOC,BTLA,ESR2,LATS1,NEDD4,SSPN*
Cellular Growth and Proliferation	1.71E-03-4.46E-02	*BOC,BTLA,ESR2,LATS1,NEDD4*
Post-Translational Modification	1.71E-03-8.52E-03	*PCMT1,SHROOM3*
Cell Death and Survival	2.69E-03-4.35E-02	*ATG3,BTLA,CCDC80,CD2AP,ESR2,FGF12, ITPR2,LATS1*
Cell Cycle	3.42E-03-4.35E-02	*BTLA,CD2AP,DYNC1LI1,ESR2,LATS1,SPICE1*
DNA Replication, Recombination, and Repair	3.42E-03-4.19E-02	*LATS1*
Cell Signaling	5.12E-03-3.27E-02	*ESR2,ITPR2,NEDD4*
Cellular Compromise	5.12E-03-4.19E-02	*DAB1,ESR2,KATNA1*
Drug Metabolism	5.12E-03-2.7E-02	*ESR2*
Lipid Metabolism	5.12E-03-4.31E-02	*ADGRF5,DAB1,ESR2,SSPN*
Protein Trafficking	5.12E-03-5.12E-03	*ESR2,SHROOM3*
Vitamin and Mineral Metabolism	5.12E-03-3.27E-02	*ESR2,ITPR2*
Carbohydrate Metabolism	1.02E-02-1.02E-02	*HS6ST3*
Cellular Response to Therapeutics	1.02E-02-2.53E-02	*ATG3,ESR2*
Gene Expression	1.7E-02-1.7E-02	*ESR2*
Free Radical Scavenging	3.36E-02-3.36E-02	*ADGRF5*

Table 3.S3 Enriched biological functions significantly associated with DMI in a beef cattle multibreed population

Biological function	*P*-value range	Genes involved in the biological function
Lipid Metabolism	2.46E-04-2.81E-02	*ACACB,ADGRF5,ALDH3B1,AQP9,CCDC80,CHKA, CPT1A,DAB1,DDX5,HNF1A,IGHMBP2,LRP5,NOS1, PLSCR1,PVALB,SSPN*
Molecular Transport	2.46E-04-2.54E-02	*ACACB,ADGRF5,AQP9,CCDC80,CD247,CHKA,CPT1A, DAB1,DDX5,FGF12,HNF1A,IFT27,IGHMBP2,ITPR2, LRP5,NEDD4,NOS1,NPFFR2,PECAM1,PLSCR1,PVAL, TCIRG1*
Small Molecule Biochemistry	2.46E-04-2.81E-02	*ACACB,ADGRF5,ALDH3B1,AQP9,CCDC80,CHKA, CPT1A,DAB1,DDX5,HNF1A,HS6ST3,IGHMBP2,LRP5, NOS1,NPFFR2,PECAM1,PLSCR1,PVALB,SGCD,SSPN, TCIRG1*
Nucleic Acid Metabolism	2.84E-04-2.81E-02	*ACACB,CPT1A,NOS1*
Protein Synthesis	5.66E-04-1.13E-02	*ACACB,CCDC80,HNF1A,LRP5,NOS1,NPFFR2, PECAM1,SGCD*
Cell Death and Survival	6.11E-04-2.81E-02	*ATG3,BTLA,CD247,DDX5,HNF1A,ITPR2,LRP5,NOS1, PECAM1,PERP,PLSCR1,POU2F1,SGCD,SSPN*
Cellular Compromise	6.11E-04-2.54E-02	*HNF1A,LRP5,NOS1,PLSCR1,PRTG,TCIRG1*
Cell Signaling	7.06E-04-2.26E-02	*BICD1,BOC,CD247,DDX5,HHAT,HNF1A,IFT27,ITPR2, LRP5,NOS1,OASL,PLSCR1,PVALB,SGCD,TCIRG1*
Cellular Function and Maintenance	1.19E-03-2.54E-02	*ATG3,BICD1,BOC,CACNG2,DAB1,FCHO2,HNF1A, IFT27,ITPR2,LRP5,NEDD4,NOS1,PECAM1,PLSCR1, PVALB,SGCD,TCIRG1*
Cell Cycle	1.62E-03-2.26E-02	*BTLA,CHKA,PLSCR1,POU2F1*
Carbohydrate Metabolism	2.38E-03-2.54E-02	*ACACB,AQP9,CCDC80,CHKA,CPT1A,HNF1A,HS6ST3, LRP5,PLSCR1*
Amino Acid Metabolism	2.85E-03-2.54E-02	*HNF1A,NOS1*
Cell Morphology	2.85E-03-2.81E-02	*ADGRF5,ATG3,BOC,CACNG2,CD247,FOXN4,HNF1A, LRP5,NEDD4,NOS1,PECAM1,PERP,POLG2,PVALB, SGCD,TCIRG1*
Cell-To-Cell Signaling and Interaction	2.85E-03-2.65E-02	*ACACB,CACNG2,CD247,CDH8,FGF12,LRP5,NEDD4, NOS1,PECAM1,PERP,PVALB*
Cellular Assembly and Organization	2.85E-03-2.81E-02	*ACACB,ATG3,BOC,CACNG2,CD247,CDH8,DAB1, FCHO2,HNF1A,IFT27,ITPR2,KMT5B,NEDD4,NOS1, PECAM1,PERP,PLSCR1,POLG2,PVALB,SGCD*
Cellular Development	2.85E-03-2.81E-02	*BOC,BTLA,CHKA,DAB1,FOXN4,HNF1A,ITPR2,NEPRO, NOS1,PECAM1,PLSCR1,PRTG,PVALB,SSPN,TCIRG1*
Cellular Growth and Proliferation	2.85E-03-2.81E-02	*BOC,BTLA,HNF1A,ITPR2,NOS1,PECAM1,PLSCR1, POU2F1,PVALB,TCIRG1*
Cellular Movement	2.85E-03-2.56E-02	*DAB1,HNF1A,MACF1,NEDD4,NOS1,PECAM1*
Drug Metabolism	2.85E-03-2.85E-03	*PVALB*

RNA Post-Transcriptional Modification	2.85E-03-2.85E-03	*DDX5*
Energy Production	5.69E-03-1.98E-02	*ACACB,CCDC80,CPT1A,DDX5,HNF1A,NOS1*
Gene Expression	5.69E-03-2.26E-02	*DDX5,IGHMBP2,POU2F1*
Post-Translational Modification	5.69E-03-8.52E-03	*HNF1A,NOS1*
Vitamin and Mineral Metabolism	8.52E-03-2.26E-02	*ACACB,CD247,HNF1A,ITPR2,LRP5,NOS1,PVALB, SGCD,TCIRG1*
Cellular Response to Therapeutics	1.7E-02-2.54E-02	*ATG3*

Table 3.S4 Enriched biological functions significantly associated with ADG in a beef cattle multibreed population

Biological function	*P*-value range	Genes involved in the biological function
Cell-To-Cell Signaling and Interaction	6.57E-04-6.57E-04	*TRAF3*
Cellular Development	6.57E-04-2.34E-02	*TRAF3*
Cellular Function and Maintenance	6.57E-04-2.21E-02	*AMN,ARL8B,MYO19,TRAF3*
Cellular Growth and Proliferation	6.57E-04-2.34E-02	*TRAF3*
Molecular Transport	6.57E-04-2.34E-02	*AMN,ARL8B,CA4,SLC28A3,TRAF3*
Nucleic Acid Metabolism	6.57E-04-1.5E-02	*SLC28A3*
Small Molecule Biochemistry	6.57E-04-1.5E-02	*AMN,SLC28A3*
Cellular Compromise	1.31E-03-1.18E-02	*ARL8B,CA4*
Cellular Movement	1.31E-03-2.15E-02	*ARL8B,CA4*
Protein Trafficking	2.63E-03-4.59E-03	*AMN,TRAF3*
Cell Death and Survival	3.28E-03-3.24E-02	*ARL8B,CA4,TRAF3*
Drug Metabolism	4.59E-03-4.59E-03	*AMN*
Cell Cycle	6.56E-03-3.81E-02	*MYO19*
Cellular Assembly and Organization	7.21E-03-3.62E-02	*ARL8B,MYO19*
RNA Post-Transcriptional Modification	8.51E-03-8.51E-03	*INTS6*
Cell Morphology	9.82E-03-9.82E-03	*ARL8B*
Post-Translational Modification	1.02E-02-1.02E-02	*TRAF3,USP32*
Vitamin and Mineral Metabolism	1.5E-02-1.5E-02	*AMN*

Table 3.S5 Enriched biological functions significantly associated with MWT in a beef cattle multibreed population

Biological function	*P*-value range	Genes involved in the biological function
Molecular Transport	7.04E-04-4.79E-02	*AMN,AQP9,ARL8B,ATP6V0A4,CA4,CAMK2B, DDX56,DDX6,NPC1L1,PPP3CC,SLC28A3, SORBS3,TRAF3, TRAK1*
Nucleic Acid Metabolism	7.04E-04-2.7E-02	*AQP9,SLC28A3*
Small Molecule Biochemistry	7.04E-04-4.79E-02	*AMN,AQP9,NPC1L1,PPP3CC,SLC28A3, SORBS3, TREH*
Cell Cycle	1.45E-03-2.78E-02	*ARL8B,CAMK2B,MYO19,NUDCD3,TRIM24*
Cell Morphology	1.45E-03-3.36E-02	*AQP9,ARL8B,CAMK2B,NUDCD3,PDLIM2, PINX1, TRIM24*
Cellular Assembly and Organization	1.45E-03-4.38E-02	*ARL8B,CAMK2B,CCAR2,CDC42BPB,DDX6, MYO16, NPC1L1,NUDCD3,PDLIM2,PHLDB1,PINX1, TRAK1, TRIM24,ULK4*
DNA Replication, Recombination, and Repair	1.45E-03-4.19E-02	*ARL8B,NUDCD3,PINX1,TRIM24*
Cellular Function and Maintenance	1.49E-03-4.38E-02	*AMN,AQP9,ARL8B,CAMK2B,CCAR2, CDC42BPB,MYO16,MYO19,NPC1L1,NUDCD3, PHLDB1,TRAF3, TRAK1,ULK4*
Carbohydrate Metabolism	1.71E-03-1.36E-02	*AQP9,TREH*
Cell-To-Cell Signaling and Interaction	1.71E-03-4.19E-02	*CAMK2B,PDLIM2,PPP3CC,SORBS3,TRAF3*
Cellular Development	1.71E-03-3.2E-02	*CAMK2B,TRAF3,ULK4*
Cellular Growth and Proliferation	1.71E-03-3.2E-02	*CCAR2,DDX56,SOX7,TRAF3*
Lipid Metabolism	1.71E-03-4.79E-02	*AQP9,NPC1L1*
Vitamin and Mineral Metabolism	1.71E-03-3.86E-02	*AMN,NPC1L1*
Cellular Compromise	3.42E-03-3.03E-02	*ARL8B,CA4,CCAR2,PINX1*
Cellular Movement	3.42E-03-3.53E-02	*ARL8B,CA4,PPP3CC,TRAK1*
Gene Expression	3.42E-03-4.62E-02	*DDX6,EIF5,TRIM24*
RNA Damage and Repair	3.42E-03-3.42E-03	*EIF5*
Post-Translational Modification	4.42E-03-3.53E-02	*CAMK2B,CDC42BPB,KLHL31,MARK3,TRIM24, UBE2E2*
Protein Trafficking	6.82E-03-1.53E-02	*AMN,TRAF3,UBE2E2*
Cell Death and Survival	8.52E-03-3.03E-02	*ARL8B,CAMK2B,PPP3CC,TRAF3*
Protein Degradation	9.57E-03-2.09E-02	*CCAR2,EDEM1,PINX1,TRAF3,TRIM24*
Protein Synthesis	9.57E-03-3.7E-02	*CAMK2B,CCAR2,DDX6,EDEM1,EIF5,MARK3, NPC1L1,PINX1,TRAF3,TRIM24*
Drug Metabolism	1.19E-02-1.19E-02	*AMN*
Protein Folding	1.53E-02-1.53E-02	*UBE2E2*
Cell Signaling	1.88E-02-3.89E-02	*CAMK2B,KLHL31,MARK3,SORBS3,ULK4*
RNA Post-Transcriptional Modification	2.2E-02-2.2E-02	*INTS6*
RNA Trafficking	3.7E-02-3.7E-02	*DDX6*

Table 3.S6 Uniquely common biological functions for feed efficiency traits in a beef cattle multibreed population

Trait	Number of common biological functions	Common biological functions
ADG, DMI, MWT, RFI	15	Molecular Transport, Cellular Function and Maintenance, Cell Morphology, Small Molecule Biochemistry, Cellular Assembly and Organization, Cell-To-Cell Signaling and Interaction, Cell Cycle, Cellular Development, Cellular Movement, Drug Metabolism, Vitamin and Mineral Metabolism, Cell Death and Survival, Cellular Growth and Proliferation, Post-Translational Modification, Cellular Compromise
DMI, MWT, RFI	5	Lipid Metabolism, Carbohydrate Metabolism, Cell Signaling, Protein Synthesis, Gene Expression
ADG, DMI, MWT	2	Nucleic Acid Metabolism, RNA Post-Transcriptional Modification
ADG, MWT, RFI	1	Protein Trafficking
DMI, RFI	1	Cellular Response to Therapeutics
MWT, RFI	1	DNA Replication, Recombination, and Repair

Table 3.S7 The comparison of candidate genes between the current study and Zhang et al.

Trait	Number of candidate genes identified by Zhang et al.	Number of candidate genes identified in this study	Number of overlapping genes	overlapping genes / total identified genes (%)	Overlapping genes
RFI	596	40	10	25.00%	*SHROOM3, GTPBP8, ADGRF5, SLC9A9, NEPRO, ADGRF1, SYNE2, FRMD5, ADGRF4, ITPR2*
DMI	268	66	23	34.85%	*GTPBP8, CUX2, ADGRF5, DDX5, NEPRO, NOS1, ADGRF1, NDUFS8, MFHAS1, FBXO21, TCIRG1, HNF1A, CHKA, C17H12orf43, LRP5, OASL, PPP6R3, TMEM171, PERP, FCHO2, POU2F1, MACF1, SGCD*
ADG	179	15	6	40.00%	*SERPINE3, INTS6, TRAF3, AMN, CDC42BPB, SLC28A3*
MWT	532	40	7	17.50%	*UBE2E2, MARK3, USP32, CAMK2B, SLC28A3, PPP3CC, SORBS3*

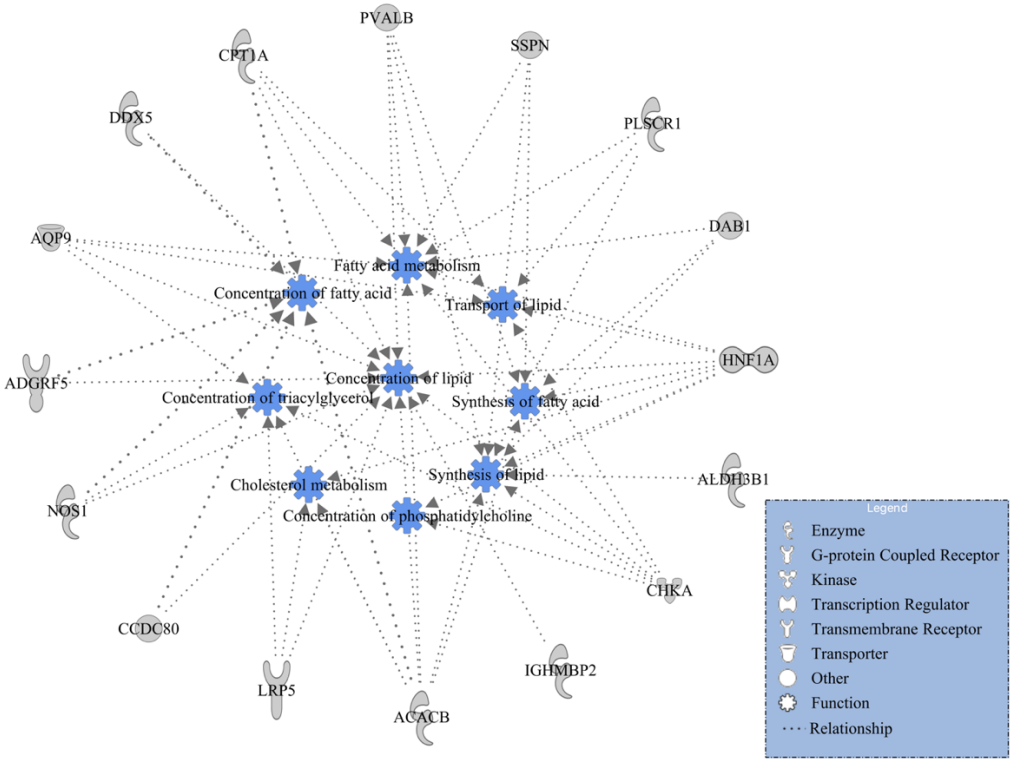

Figure 3.1 Gene network of lipid metabolism for dry matter intake (DMI)

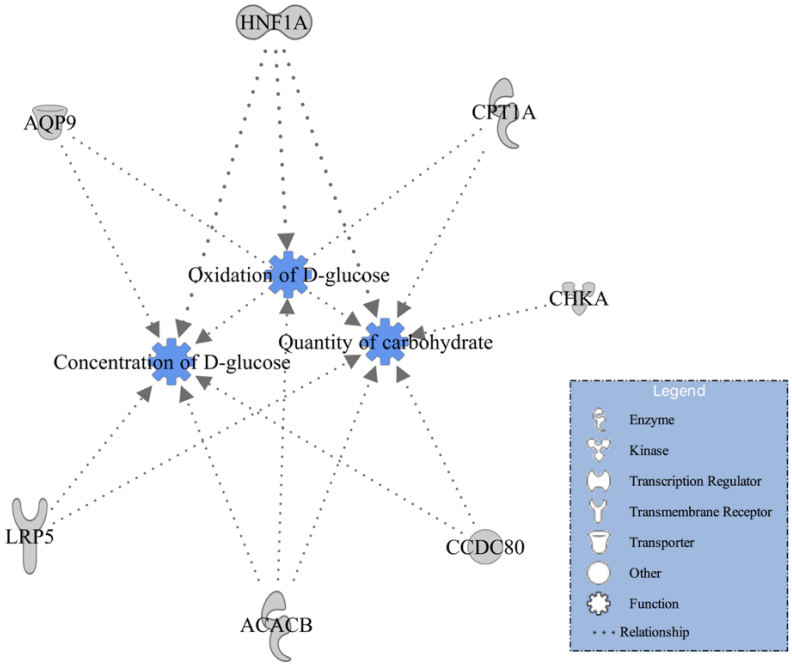

Figure 3.2 Gene network of carbohydrate metabolism for dry matter intake (DMI)

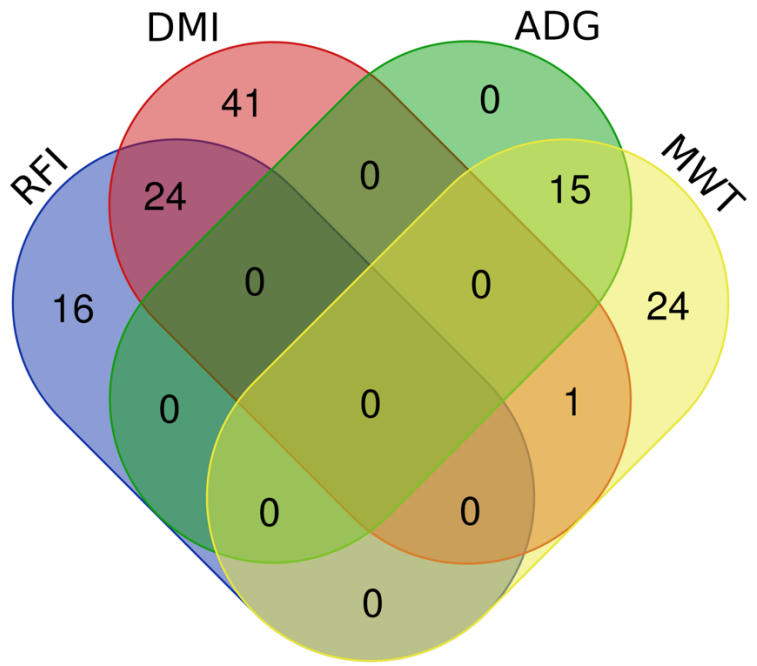

Figure 3.S1 Uniquely common candidate genes for feed efficiency traits RFI, DMI, ADG, MWT in a beef cattle multibreed population

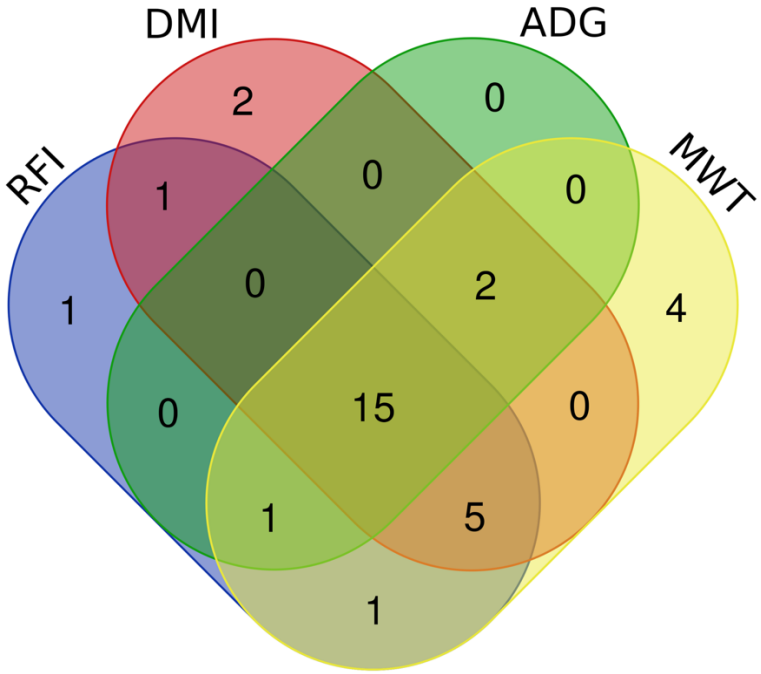

Figure 3.S2 Uniquely common biological functions for feed efficiency traits RFI, DMI, ADG, MWT in a beef cattle multibreed population

Chapter 4. Integrative analyses of genomic and metabolomic data reveal genetic mechanisms associated with carcass merit traits in beef cattle

4.1 Abstract

Improvement of carcass merit traits is a priority for the beef industry. Discovering DNA variants and genes associated with variation in these traits and understanding biological functions/processes underlying their associations are of paramount importance for more effective genetic improvement of carcass merit traits in beef cattle. This study integrates 10,488,742 imputed whole genome DNA variants, 31 plasma metabolites, and animal phenotypes to identify genes and biological functions/processes that are associated with carcass merit traits including hot carcass weight (HCW), rib eye area (REA), average backfat thickness (AFAT), lean meat yield (LMY), and carcass marbling score (CMAR) in a population of 493 crossbred beef cattle.

Regression analyses were performed to identify plasma metabolites associated with the carcass merit traits, and the results showed that 4 (3-hydroxybutyric acid, acetic acid, citric acid, and choline), 6 (creatinine L-glutamine, succinic acid, pyruvic acid, L-lactic acid, and 3-hydroxybutyric acid), 4 (fumaric acid, methanol, D-glucose, and glycerol), 2 (L-lactic acid and creatinine), and 5 (succinic acid, fumaric acid, lysine, glycine, and choline) plasma metabolites were significantly associated with HCW, REA, AFAT, LMY, and CMAR (P-value < 0.1), respectively. Combining the results of metabolome-genome wide association studies using the 10,488,742 imputed SNPs, 103, 160, 83, 43, and 109 candidate genes were identified as significantly associated with HCW, REA, AFAT, LMY, and CMAR (P-value < 1×10^{-5}), respectively. By applying functional enrichment analyses for candidate genes of each trait, 26, 24, 26, 24, and 28 significant cellular and molecular functions were predicted for HCW, REA, AFAT, LMY, and CMAR, respectively. Among the five topmost significantly enriched biological

functions for carcass merit traits, molecular transport and small molecule biochemistry were two top biological functions associated with all carcass merit traits. Lipid metabolism was the most significant biological function for LMY and CMAR and it was also the second and fourth highest biological function for REA and HCW, respectively. Candidate genes and enriched biological functions identified by the integrative analyses of metabolites with phenotypic traits and DNA variants could help to interpret the results of previous genome-wide association studies for carcass merit traits. Our integrative study also revealed additional potential novel genes associated with these economically important traits.

Therefore, our study improves understanding of the molecular and biological functions/processes that influence carcass merit traits, which could help develop strategies to enhance genomic prediction of carcass merit traits with incorporation of metabolomic data. Similarly, this information could guide management practices, such as nutritional interventions, with the purpose of boosting specific carcass merit traits.

4.2 Introduction

Carcass merit traits, including hot carcass weight (HCW), rib eye area (REA), average backfat thickness (AFAT), lean meat yield (LMY), and carcass marbling score (CMAR), are economically important traits in beef cattle since they directly influence the meat product yield and quality grade, and therefore profitability. For example, sufficient marbling is important for beef tenderness, juiciness and flavor, so the degree of marbling in beef is the primary factor determining quality grade of the meat. However, carcass merit traits are expressed late in life and the measurement of these traits for individual live animals is relatively expensive via ultrasound technologies. In many cases evaluation occurs post mortem, thereby eliminating breeding stock with superior breeding values for the traits. The development of genomic prediction provides an

opportunity to assess genetic merit of animals as early as birth (Todd et al., 2014; Fernandes Júnior et al., 2016; Mehrban et al., 2017; Xu et al., 2020) but there is still a need to improve the accuracy of genomic selection for carcass traits in beef cattle in order to achieve broader industry applications (Chen et al., 2015; Mehrban et al., 2017; Ogawa et al., 2017). Detecting more candidate genes and potentially functional or causal DNA variants through genome-wide association studies (GWAS) and understanding the biological background of the relationship between the genome and phenome could help improve the accuracy of genomic selection for complex traits including carcass merit traits (Meuwissen et al., 2013, 2021; Snelling et al., 2013; Zhang et al., 2019).

As more omics-based intermediate phenotypes, based on gene expression, protein and metabolite analysis, become available, integrating multi-omics data to further elucidate genetic influence of complex traits holds great promise (Widmann et al., 2013; Weber et al., 2016; Fonseca et al., 2018, 2019). Among the omics-based intermediate phenotypes, metabolites have been reported to be associated with carcass merit traits of livestock (Matthews et al., 2001; Connolly et al., 2019; Goldansaz et al., 2020) and their variation is influenced by genetic effects (Buitenhuis et al., 2013; Li et al., 2020). Therefore, we hypothesize that combining metabolomic data into GWAS of whole genome DNA variants could help detect key candidate genes and potentially functional or causal DNA variants associated with carcass merit traits.

In this study, the data of 5 carcass merit traits (HCW, REA, AFAT, LMY, and CMAR) and 31 plasma metabolites were collected from a beef cattle population consisting of 493 crossbred bulls, heifers, and steers. Our objective was to identify significant single nucleotide polymorphisms (SNPs), candidate genes and biological functions associated with carcass merit traits through integration of carcass merit traits, plasma metabolites and whole genome sequence

variants. Linear regression models were first used to identify metabolites associated with carcass merit traits. Metabolome-genome wide association studies (mGWAS) were then performed with 10,488,742 imputed whole genome SNPs to identify significant SNPs for the trait associated metabolites. Candidate genes were mapped based on significant SNPs and gene functional enrichment analyses were subsequently performed on candidate genes of each trait to predict biological functions/processes associated with carcass merit traits in beef cattle.

4.3 Material and Methods

4.3.1 Animal population, metabolomic and phenotypic data collection

All animals in this study were cared for according to the guidelines of the Canadian Council on Animal Care (2009) and the experimental procedures were approved by the University of Alberta Animal Care and Use Committee (AUP00000777). In total, 493 bulls (n = 93), heifers (n = 125) or steers (n = 275) from different herds including Charolais (n = 73), Hereford-Angus crosses (n = 191), and Beefbooster composite breed (predominantly Charolais-based, n = 229) were used. Among these animals, 277 animals from two herds had implants, while 216 animals from the other three herds had no implant. The effect of the factor of implant was examined using statistical analysis, and its effect has been captured under the herd variable applied in subsequent statistical analysis. Animals were born between 2002 to 2011 and initially measured for feed intake using the GrowSafe system (GrowSafe Systems Ltd., Airdrie, Alberta, Canada) at the feedlot test station under multiple projects, which were described previously (Basarab et al., 2003; Nkrumah et al., 2007b; Basarab et al., 2011; Zhang et al., 2020). The animals from the same herd of a particular year were tested in the same feedlot and diet. Blood samples were collected from all animals by jugular venipuncture in the early morning on the first day of feedlot feeding test and immediately frozen at -80°C for storage. Plasma metabolites were quantified using nuclear

magnetic resonance (NMR) spectroscopy as described by Li et al. (2020). Briefly, blood samples were thawed at room temperature and centrifuged at 10,000 rpm for 10 minutes to separate the plasma. Plasma was then filtered through 3kDa molecular weight cut-off filters (Merck Millipore Ltd. 3KDA filter tubes; Darmstadt, Germany) to exclude macromolecules, including lipids and proteins. After filtration samples with a low volume were diluted with high-performance liquid chromatography (HPLC) water to 570 µl to ensure an adequate volume for NMR acquisition. Samples were further prepared in a 5 mm NMR tube (New Era Enterprises Inc., NJ, USA) that contained a total of 700 µl including 570 µl filtered serum, 60 µl DSS and 70 µl D2O. Spectra were acquired on a 500MHz VNMRS spectrometer equipped with a 5mm cold probe (Agilent Technologies, CA, USA). Metabolites were identified and quantified with a targeted profiling approach using the Profiler and Library Manager modules in the same software which contained 304 total metabolites as references. Each spectrum was reviewed by a separate analyst and a final review was performed on all of the spectra before exporting concentration results. Concentration measurements were adjusted to report metabolite concentrations (µM). In total, 33 metabolites were initially identified and quantified using NMR. However, two of them were excluded due to missing values, resulting in 31 metabolites for further analyses.

In order to collect carcass data, animals were then slaughtered after the feedlot tests at either a commercial plant or the Lacombe Research and Development Centre (LRDC) abattoir when a majority of them reached > 8 mm backfat as predicted from ultrasound measurement. The processes of carcass data collection were previously described (Nkrumah et al., 2007a; Basarab et al., 2011; Lu et al., 2013; Mao et al., 2013; Akanno et al., 2014; Chen et al., 2014). In summary, hot carcass weight (HCW) in kg was obtained by summing up the weight of each side of the carcass that was split during dressing, about 45 min post-mortem. Average backfat thickness (AFAT) in

mm, rib eye area (REA) in square centimeters, and carcass marbling score (CMAR) at the grading site between the 12th and 13th ribs was assessed by trained personnel. Carcass marbling score was measured as a continuous variable from 100 (trace marbling or less) to 499 (abundant or more marbling) to reflect the amount of fat deposit interspersed between the muscle fibers (i.e., intramuscular fat) of the longissimus thoracis. Lean meat yield (LMY) was calculated as LMY, % = 57.96 + (0.202 × REA, cm^2) − (0.027 × HCW, kg) − (0.703 × AFAT, mm) as an estimate of saleable meat in the carcass (Basarab et al., 2003).

4.3.2 Animal genotyping, SNP imputation and quality control

DNA was extracted from the blood samples using DNeasy Blood & Tissue Kit (QIAGEN, Ontario, Canada). and then genotyped using the Illumina BovineSNP50 v2 BeadChip (Illumina Inc., CA, USA). For the SNP imputation, a step-wise procedure was applied using Beagle 5.1 software (Browning et al., 2018) as described by Zhang et al. (2020) and Wang et al. (2020). Briefly, we first imputed from the 50K SNPs to the AffyHD panel of 444,558 SNPs with 4,247 animals of mixed beef breeds in the reference population. We then imputed from the imputed AffyHD panel to the whole genome sequence variants with the reference population of 3,093 Bos taurus animals from the 1000 Bull Genomes Project (Daetwyler et al., 2014) (run 7). Finally, 53,258,178 variants (SNPs and indels) on 29 autosomes were obtained after the imputation with average imputation accuracy of 0.97 across variants with a standard deviation (SD) of 0.08, which was assessed through a five-fold cross-validation as described by Zhang et al. (2020) and Wang et al. (2020). For each SNP, post-imputation quality control was then performed to filter the imputed variant genotypes if one of the following conditions was met: (1): SNPs on 29 autosomes that had an imputation accuracy < 0.95; (2): minor allele frequency < 0.05; (3) SNPs failed to pass the

Hardy-Weinberg equilibrium test (P-value < 0.0001). A total of 10,488,742 SNPs remained for subsequent analyses after the quality control.

4.3.3 Regression analyses between carcass merit traits and metabolites and metabolome-genome wide association studies

Phenotypic values of carcass merit traits and metabolites were corrected for factors including animal type (bull, heifer, steer), birth year, herd, feedlot pen, animal age at slaughter, and breeding composition using linear regression models. The breed composition ($K = 6$) of each animal was predicted based on their 50K SNPs using ADMIXTURE software to account for population stratification (Alexander et al., 2009; Hellwege et al., 2017). Residuals of metabolomic and phenotypic data beyond 3 standard deviations from the mean of residuals were considered as outliers and excluded from further analyses. In order to determine relationships between carcass merit traits and metabolites, regression analyses were conducted between the adjusted carcass merit traits and the 31 adjusted metabolites using R statistical software. A carcass merit trait and a metabolite were considered to be significantly associated when the regression analyses have *P*-value < 0.1.

For the metabolome-genome wide association studies (mGWAS), the adjusted values of metabolites that were significantly associated with the carcass merit traits were the response variable in the single SNP-based mixed linear model association (*mlma*) as implemented in GCTA software (Yang et al., 2011). The linear mixed model can be described as follows:

$$y_{ij} = \mu + b_j x_{ij} + a_{ij} + e_{ij}$$

where y_{ij} is the adjusted metabolite value of the ith animal with the jth SNP (i.e. the ijth animal), b_j is the allele substitution effect of the jth SNP, x_{ij} is the jth SNP genotype of animal i coded as 0, 1, 2 for genotypes A_1A_1, A_1A_2, and A_2A_2, respectively, a_{ij} is the additive polygenic

effect of the ijth animal $\sim N(0, \boldsymbol{G}\sigma_a^2)$, and e_{ij} is the random residual effect $\sim N(0, \boldsymbol{I}\sigma_e^2)$. The genomic relationship matrix $\boldsymbol{G}$ that was derived based on total filtered SNP markers (i.e. 10,488,742 SNPs) as described by Yang et al. (2014), which is essentially the same as the second VanRaden's formulation (VanRaden, 2008). The SNP allele substitution effect was estimated and the significance test of the SNP allele substitution effect was conducted via a generalized least square F-test as implemented in the GCTA package. The SNPs with *P*-value < 1×10^{-5} were considered to be significantly associated with the metabolite according to the recommendation of The Wellcome Trust Case Control Consortium (Burton et al., 2007). The phenotypic variance of the metabolite explained by each significant SNP was calculated by $\frac{2pq\beta^2}{S^2} * 100\%$, where p and q denote the SNP allele frequency of A_1 and A_2, respectively; β is the SNP allele substitution effect; $2pq\beta^2$ is the additive variance of the SNP, and S^2 is the phenotypic variance of the metabolite.

4.3.4 Identification of candidate gene and functional enrichment analyses for carcass merit traits

A 140-kbp window (70-kbp upstream and 70-kbp downstream) of each significant SNP was used to map candidate genes based on ARS-UCD 1.2 bovine genome assembly from the Ensembl BioMart database (accessed in February 2021). The 70-kbp was the chromosomal length within which a high linkage disequilibrium phase correlation ($r^2 > 0.77$) was maintained across a sample of Canadian beef cattle breeds (Lu et al., 2012).

The Entrez gene IDs were used as gene identifiers and small nucleolar RNA and microRNA were excluded from gene functional enrichment analyses. Bovine genes were changed to known human orthologous genes from Ensembl, whereas for those genes without human orthologs their bovine gene IDs were maintained in the gene list. Then candidate genes of all metabolites associated with the carcass merit traits (HCW, REA, AFAT, LMY or CMAR) as

identified in the regression analyses between carcass merit traits and metabolites were combined and imported into the Ingenuity Pathway Analysis software (accessed in February 2021) (IPA; www.Ingenuity.com). Significantly enriched molecular and cellular biological functions and gene networks (*P*-value < 0.05) for each carcass merit traits were inferred and gene-sub-biological function/process interactions within the most significant molecular and cellular functions were predicted in the IPA.

4.4 Results

4.4.1 Metabolites associated with carcass merit traits

The results of regression analyses showed 15 out of 31 analyzed metabolites were associated with one or more than one of the carcass merit traits (Table 4.1). At *P*-values less than 0.05, 3 (3-hydroxybutyric acid, acetic acid, and citric acid), 3 (creatinine, L-glutamine, and succinic acid), 1 (methanol), 1 (L-lactic acid), and 3 (succinic acid, lysine, and glycine) metabolites were identified as associated with HCW, REA, AFAT, LMY, and CMAR, respectively. Some of metabolites (*P*-value from 0.05 to 0.1) explained more than 1% of the phenotypic variance of associated carcass merit traits, thus, a relatively relaxed threshold of *P*-value < 0.1 was chosen to include more metabolites that may be potentially associated with carcass merit traits. For HCW, at *P*-values less than 0.1, 3-hydroxybutyric acid, acetic acid, citric acid and choline were the associated metabolites, accounting for 1.92%, 1.69%, 1.48%, and 1.09% of the phenotypic variance in HCW, respectively. Creatinine, L-glutamine, succinic acid, pyruvic acid, L-lactic acid and 3-hydroxybutyric acid were significantly associated with REA, and these six metabolites accounted for 1.87%, 1.79%, 1.60%, 1.56%, 1.04%, and 0.80% of phenotypic variance, respectively. AFAT was associated with fumaric acid, methanol, D-glucose and glycerol, and these four metabolites explained 2.40%, 1.71%, 1.67%, and 1.39% of the phenotypic variance,

respectively. L-lactic acid and creatinine were associated with LMY and accounted for 2.71% and 1.42% of phenotypic variance, respectively. Five metabolites, including succinic acid, fumaric acid, lysine, glycine, and choline, were associated with CMAR and each respectively accounted 1.76%, 1.36%, 1.22%, 1.20%, and 1.05% of the phenotypic variance in CMAR, respectively. Most of these metabolites were mainly associated with a single trait. However, a few metabolites were associated with more than one trait. For example, 3-hydroxybutyric acid was associated with both HCW and REA. L-lactic acid and creatinine were both associated with REA and LMY. Choline was associated with HCW and CMAR, and fumaric acid was associated with both AFAT and CMAR (Table 4.1).

4.4.2 Significant SNPs and candidate genes associated with metabolites

Genomic inflation factors (defined as the median of the observed chi-squared test statistics divided by the expected median of the corresponding chi-squared distribution) for all association analyses ranged from 0.95 to 1.01, the value around 1 indicates that there is no population stratification, and the statistical models are well fitted. Summarized results of the mGWAS for the 15 metabolites (identified as associated with the carcass merit traits, *P*-value < 0.1) are presented in Table 4.2. The average of phenotypic variance of the metabolites explained by a single SNP was 5.13% with a range of 3.57% - 10.95%. Through integrating the metabolite and carcass merit trait regression analyses and the mGWAS results, a total of 103, 160, 83, 43, and 109 candidate genes were found to be associated with HCW, REA, AFAT, LMY, and CMAR, respectively (Table 4.3). As for metabolites, some candidate genes identified through the mGWAS were associated with multiple carcass traits (Table 4.S1 and Figure 4.S1). For instance, *CDH13* was associated with HCW, REA, AFAT, and CMAR, while 5 genes (*KMT5B*, *NDUFS8*, *ALDH3B1*, *CHKA*, and *TCIRG1*) were associated with HCW, REA, LMY, and CMAR.

4.4.3 Significantly enriched molecular functions and gene networks for carcass merit traits

Of the identified candidate genes for HCW, REA, AFAT, LMY and CMAR, 99, 149, 78, 42 and 102 genes were respectively mapped to the IPA database for functional enrichment analyses. Briefly, 26, 24, 26, 24, and 28 cellular and molecular functions were identified as significantly (*P*-value < 0.05) associated with HCW, REA, AFAT, LMY, and CMAR, respectively (Table 4.S2 - 4.S6). Interestingly, 75% of the biological functions were commonly associated with all the five carcass merit traits in this study (Table 4.S7 and Figure 4.S2). Some of the major functions common across the traits included molecular transport, small molecule biochemistry, lipid metabolism, cell-to-cell signaling and interaction, carbohydrate metabolism, cellular assembly and organization. Additionally, the five topmost significantly enriched biological functions for each trait and the candidate genes involved are presented in Table 4.4. Among the top five significant enriched functions, molecular transport and small molecule biochemistry were commonly associated with all carcass merit traits. Lipid metabolism was the most significant biological function for LMY and CMAR, and it was the second and fourth highest biological function for REA and HCW, respectively. Cell-to-cell signaling and interaction was one of the top significant biological functions associated with HCW, REA, AFAT, and CMAR. Carbohydrate metabolism was among the top significant biological functions associated with both HCW and CMAR. Further investigation within some of the biological functions revealed molecular/metabolic processes related to the carcass merit traits. For HCW, within the molecular transport function, 11 genes (*AGTR1*, *CHKA*, *CPT1A*, *DDX5*, *IGHMBP2*, *IL21R*, *LRP5*, *NTRK2*, *PVALB*, *SULT1E1*, and *TRAF3*) were involved in concentration of corticosterone and lipid, and quantity of steroid and steroid hormone (Figure 4.1a). For CMAR, within the lipid metabolism function, 17 candidate genes (*AGTR1*, *AQP9*, *CCDC80*, *CHKA*, *CPT1A*, *DAB1*, *DDX5*, *IGHMBP2*, *LRP5*, *PLA2G2A*,

PLA2G2E, *PLA2G5*, *PTGS1*, *PVALB*, *SLC10A1*, *SPRED2*, and *SSPN*) were involved in multiple metabolic processes related to fatty acid and lipid metabolism (Figure 4.1b), such as synthesis of fatty acid and release of lipid. Additionally, within the carbohydrate metabolism function for CMAR, 12 candidate genes (*AGTR1*, *AQP9*, *CHKA*, *CPT1A*, *GYG1*, *LRP5*, *NKX3-2*, *PDCL*, *PLA2G2A*, *PLA2G2E*, *PLA2G5*, and *TREH*) were involved in carbohydrate metabolic processes such as carbohydrate biosynthesis (Figure 4.1c). It is worth noting that 8 candidate genes (*AGTR1*, *AQP9*, *CHKA*, *CPT1A*, *LRP5*, *PLA2G2A*, *PLA2G2E*, and *PLA2G5*) associated with CMAR were involved in both lipid and carbohydrate metabolism.

4.5 Discussion

4.5.1 Metabolomics to improve understanding on genetic influence of carcass merit traits

Studies have demonstrated metabolites as potential biomarkers for economically important traits in livestock species (Matthews et al., 2001; Connolly et al., 2019; Goldansaz et al., 2020). However, improving understanding of the biology involved is hampered by the limited knowledge of how these metabolites are associated with different economically important traits in the different livestock species. Carcass traits are of fundamental interest to every beef producer and everyone involved in the beef industry. However, these traits are relatively expensive to measure using ultrasound technologies on individual live animals, which is a limitation for selection and improvement of these traits. Since blood metabolites are easily measurable/quantifiable even on live animals, we speculate that identification of genetic/biological associations between metabolite concentrations and beef cattle carcass merit traits could potentially enhance genetic prediction and selection for these traits in beef cattle. In addition, identification of blood metabolite biomarkers associated with carcass traits at an earlier stage would have a more practical application for genetic selection and for sorting animals into different finishing groups for more uniform carcass outputs.

Therefore, we collected the blood samples at the start of the feedlot test instead of close to slaughter and examined the associations between 31 plasma metabolites and 5 carcass merit traits. We further explored the potential biological linkage between these metabolites and carcass merit traits. Our results showed that several metabolites were associated with the carcass merit traits studied. However, individual metabolites, despite being significantly associated with the trait, only accounted for 0.80% to 2.71% of the total phenotypic variance of carcass merit traits. This relatively small percentage of phenotypic variance reflected the complex nature of these traits, which we believe are regulated by multiple metabolic pathways involving many metabolites with each having only a small contribution/effect. It is also possible that due to the limited number of metabolites we profiled in the current study, we were not able to identify those metabolites with major influences on the traits studied. Additionally, this study used a more relaxed threshold (*P*-value < 0.1) to identify metabolites potentially associated with the carcass merit traits, therefore, the validation in independent beef cattle populations or further studies considering a wider range of metabolites is warranted. It is also worthwhile to analyze metabolites on samples collected at different developmental stages to see whether and how the associations between the metabolites and the carcass traits may change. Furthermore, we observed that a majority of the significant metabolites were only associated with one trait. However, some metabolites in the current study were associated with two traits, indicating potential biological relationships between these traits. For example, in this study, we observed that 3-hydroxybutyric acid was associated with both HCW and REA, and beef cattle with high HCW and REA had lower concentration of 3-hydroxybutyric acid, indicating that animals with high HCW and REA may have better carbohydrate metabolism. Additionally, 3-hydroxybutyrate is the main representative of ketone bodies and one important function of ketone bodies is to provide acetoacetyl-CoA and acetyl-CoA for the synthesis of

cholesterol, fatty acids, and complex lipids (Mierziak et al., 2021). Thus, a lower concentration of 3-hydroxybutyric acid may lead to reduced lipid synthesis in animals with high HCW and REA. Interestingly, creatinine, the final catabolite of muscle energy metabolism (Wyss and Kaddurah-Daouk, 2000), was positively associated with both REA and LMY in the current study (Table 4.1), and these two carcass merit traits measure muscle development and the proportion of lean meat in a carcass respectively. In line with our results, Hanset and Michaux (1982) previously observed higher concentrations of plasma creatinine in double muscled bulls as compared to conventional or normal muscled bulls, and Patel et al. (2013) proposed creatinine in serum as a promising biomarker for human muscle mass. These previous studies and the results from our study demonstrate that creatinine is a potential indicator trait or biomarker for muscle related traits in beef cattle.

4.5.2 Candidate genes, enriched molecular functions and gene networks for carcass merit traits

Generally, identification of SNPs and genes associated with carcass merit traits mainly relies on association studies between DNA variants and the traits. For example, Wang et al. (2020) performed GWAS based on 7.8 million imputed whole genome sequence variants for carcass merit traits using Canadian beef cattle and they identified hundreds of candidate genes associated with carcass merit traits. However, the knowledge about the underlying biological background behind these associations is relatively limited. We assume that metabolites, which are an intermediate phenotype lying between genome and carcass merit traits, could provide additional insight into the associations. In the current study, the candidate genes identified through incorporating metabolites showed relatively good consistency with the previous study (Wang et al., 2020) (Table 4.S8). Briefly, we found that 34.95%, 28.13%, 27.71%, 41.86%, and 22.94% of the candidate genes

identified in this study overlapped with those from Wang et al. (2020) for HCW, REA, AFAT, LMY, and CMAR, respectively. Of note, some of the candidate genes were also reported in different cattle breeds or populations in other studies. For example, *ST18* was associated with AFAT in the current study and it was associated with the metabolite D-glucose. Medeiros de Oliveira Silva et al. (2017) also identified *ST18* as candidate gene for backfat through GWAS in a Nelore cattle population. Additionally, by integrating metabolomic data, this study added more information to some previously identified associations between genes and carcass merit traits. For example, Wang et al. (2020) reported that *UMODL1*, *L3HYPDH*, *JKAMP,* and *LUZP2* were candidate genes associated with REA and LMY, but the potential mechanism of how these genes could influence the two traits remained unclear. Our study showed that these same genes (*UMODL1*, *L3HYPDH*, *JKAMP*, and *LUZP2*) were associated with the concentration of creatinine which is a metabolite associated with REA and LMY. These results indicated that these genes may be associated with the synthesis or degradation of creatinine in animals and thereby influence the related traits. Similarly, *HLTF*, *GYG1*, *RYR2*, *RBM47*, and *APBB2* were reported to be associated with REA and CMAR by Wang et al. (2020). Our results showed these genes were associated with succinic acid which was negatively associated with both REA and CMAR. Both examples represent one of situations that genes may influence different traits by regulating the same metabolites, and the mechanism of how these genetic variants affect the concentration of metabolites still needs more studies. According to our results, we would like to highlight that some genes could affect the same carcass merit traits by influencing different metabolites. For instance, *AMN* was associated with both 3-hydroxybutyric acid and citric acid, and both metabolites were identified as associated with HCW. Therefore, information obtained via analyzing metabolites could improve the understanding of genetic effects on these phenotypes. In our companion paper

by Li et al. (2021), similar findings were also observed. These two studies indicate that metabolites play important roles in the variation of both feed efficiency and carcass merit traits. Integration of metabolomic and genomic data could help to identify functional or causal SNPs or genes, and interpret the biological meaning of the candidate genes identified in GWAS. In addition, these two studies investigated the association between different omics levels, which shed light on the interrelationship between different omics layers and potential molecular mechanisms underlying these traits. Therefore, our findings have broadened our knowledge on the genetic and molecular mechanisms of these traits. Based on what we learned from these two studies, we recommend applying such multi-omics analysis to study other important traits in beef cattle.

In addition to adding more information to known associations, incorporating metabolomic data can help us identify additional novel associations as metabolites represent a level close to the final phenotypes (i.e., carcass merit traits). In the current study, some additional candidate genes were reported to be associated with carcass merit traits. Therefore, we expect that including the candidate gene SNPs in the DNA marker panel or increasing the weight applied to such SNPs could either improve accuracy of genomic prediction of the traits or decrease the DNA marker density used in genomic prediction while retaining accuracy. A preliminary attempt of this latter option was done by Melzer et al. (2013) for the prediction of three traditional milk traits in dairy cows. Melzer et al. (2013) applied regression methods to identify important milk metabolites and then those SNPs with significant genetic effects on important metabolites were identified and used to predict milk traits. Compared with the classical approach that uses all SNPs (40,317) in prediction, this metabolite approach could achieve similar prediction precision with less than 1% of the total amount of SNPs. Fontanesi (2016) suggested integration of metabolomic data would be useful if the heritability of a trait is low, if a trait is hard to be precisely and directly measured

on the animals, or if the prediction accuracy was limited by the small number of phenotyped animals in the training populations. Since carcass merit traits are expressed at later stages of animal production and are usually measured by sacrificing potential breeding stock, these traits are more suitable for DNA marker based genomic prediction, and incorporating metabolomic data into the genomic prediction has the potential to enhance the genomic prediction accuracy. In addition to metabolites, the information carried by other omics data, such as RNA and protein, also helps to prioritize SNPs associated with complex traits, and can further contribute to improving genomic prediction accuracy of these traits. For example, Fang et al. (2017) applied an extended genomic best linear unbiased prediction (GBLUP) model called genomic feature BLUP (GFBLUP) that includes a separate random effect for the joint action of SNPs within genomic features which are obtained from RNA differential expression analyses. Compared to GBLUP, the accuracy of genomic prediction for mastitis and milk production traits with GFBLUP was marginally improved (3.2 to 3.9%) in within-breed prediction but significantly increased (164.4%) in across-breed prediction. Theoretically, the genomic features could be defined from various sources of biological knowledge (e.g., metabolomics) and the GFBLUP model could be applied to other complex traits. Therefore, it would be of interest to investigate the accuracy of prediction for carcass merit traits with and without utilizing multi-omics data.

In order to further investigate biological functions associated with carcass merit traits, the five topmost significant biological functions associated with each trait were identified in the current study. These top five biological functions showed substantial overlap with the top five biological functions identified by previous studies (Wang et al., 2017, 2020; Mukiibi et al., 2018), which indicated those overlapping top biological functions potentially have important biological meaning for the carcass merit traits in beef cattle. Since our carcass data were collected from

animals that were finished for meat production, genes involved in these overlapping top biological functions, such as lipid metabolism and carbohydrate metabolism, likely play a more important role in determining the carcass merit traits. Therefore, the identification of top and other enriched biological functions and their corresponding genes will not only improve our understanding of the underlying biology but also help prioritize candidate genes and related causal SNPs for future studies. Additionally, construction of gene networks for biological functions could help us elucidate complicated connections among genes and disentangle the potential relationships among genes, biological functions and phenotypes. For example, molecular transport was identified as a top enriched biological function associated with all carcass merit traits and its network of HCW as an example showed that some of the associated genes were involved in concentration of lipid and corticosterone, and quantity of steroid and steroid hormones (Figure 4.1a). In beef cattle production, more than 30 commercially-available steroid hormone implants are marketed in the U.S. and the effects of steroid hormone implants on improving carcass leanness, increasing average daily gain, and altering dry matter intake has been reviewed by Smith and Johnson (2020). Thus, those genes linked to the functions of steroid and steroid hormones in the network may consequently influence final muscle mass in the carcass. For those genes involved in the concentration of lipid, they may influence fats in the carcass by regulating breakdown or storage of fats. Additionally, we would like to highlight the network of lipid metabolism for CMAR (Figure 4.1b) because lipid metabolism was the most significant biological function associated with this trait. In this network, some genes were involved in fatty acid metabolism including fatty acid synthesis, release and concentration. The phenotypic and genetic correlations between fatty acid composition and marbling have been reported in different beef cattle populations (Kazala et al., 1999, 2006; Hwang and Joo, 2016; Ekine-Dzivenu et al., 2017). Our results provide further

insight into the potential molecular and genetic background accounting for genetic correlations between marbling and fatty acid composition in beef cattle, further indicating that the selection for fatty acid composition or concentration could influence marbling in beef cattle as previously proposed (Ekine-Dzivenu et al., 2017).

4.6 Conclusion

In this study, genomic, metabolomic, and phenotypic data were integrated to investigate biological functions/processes related to carcass merit traits in beef cattle. Plasma metabolites associated with HCW, REA, AFAT, LMY, and CMAR were identified and individual metabolites were found to account for a small proportion of the total phenotypic variance of the carcass merit traits. Several candidate genes as associated with carcass merit traits were identified through mGWAS along with regression analyses. These genes are involved in multiple biological functions that are related to the associated carcass merit traits. Additionally, the results of our integrative analyses could help to interpret previous results from DNA marker based GWAS of the carcass merit traits and revealed functional genes associated with these economically important traits. Therefore, our integrative study has provided insights into relationships between genes, metabolites and carcass merit traits, which could potentially lead to improvement of genomic prediction accuracy via incorporating metabolomic related data.

4.7 References

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle - An independent validation study. *J. Anim. Sci.* 92, 2896–2904. doi:10.2527/jas.2013-7374.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109.

Basarab, J. A. A., Colazo, M. G. G., Ambrose, D. J. J., Novak, S., McCartney, D., and Baron, V. S. S. (2011). Residual feed intake adjusted for backfat thickness and feeding frequency is independent of fertility in beef heifers. *Can. J. Anim. Sci.* 91, 573–584. doi:10.4141/cjas2011-010.

Basarab, J. A., Price, M. A., Aalhus, J. L., Okine, E. K., Snelling, W. M., and Lyle, K. L. (2003). Residual feed intake and body composition in young growing cattle. *Can. J. Anim. Sci.* 83, 189–204. doi:10.4141/A02-065.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi:10.1016/j.ajhg.2018.07.015.

Buitenhuis, A. J. J., Sundekilde, U. K. K., Poulsen, N. A. A., Bertram, H. C. C., Larsen, L. B. B., and Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *J. Dairy Sci.* 96, 3285–3295. doi:10.3168/jds.2012-5914.

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678. doi:10.1038/nature05911.

Chen, L., Mao, F., Crews, D. H., Vinsky, M., and Li, C. (2014). Phenotypic and genetic relationships of feeding behavior with feed intake, growth performance, feed efficiency, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci.* 92, 974–983. doi:10.2527/jas.2013-6926.

Chen, L., Vinsky, M., and Li, C. (2015). Accuracy of predicting genomic breeding values for carcass merit traits in Angus and Charolais beef cattle. *Anim. Genet.* 46, 55–59. doi:10.1111/age.12238.

Connolly, S., Dona, A., Wilkinson-White, L., Hamblin, D., D'Occhio, M., and González, L. A. (2019). Relationship of the blood metabolome to subsequent carcass traits at slaughter in feedlot Wagyu crossbred steers. *Sci. Rep.* 9, 15139. doi:10.1038/s41598-019-51655-2.

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., Van Binsbergen, R., Brøndum, R. F., et al. (2014). Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* 46, 858–865. doi:10.1038/ng.3034.

Ekine-Dzivenu, C., Vinsky, M., Basarab, J. A., Aalhus, J. L., Dugan, M. E. R., and Li, C. (2017). Phenotypic and genetic correlations of fatty acid composition in subcutaneous adipose tissue with carcass merit and meat tenderness traits in Canadian beef cattle. *J. Anim. Sci.* 95, 5184–5196. doi:10.2527/jas2017.1966.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017). Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle by mapping variants to hepatic transcriptomic regions responsive to intra-mammary infection. *Genet. Sel. Evol.* 49, 44. doi:10.1186/S12711-017-0319-0.

Fernandes Júnior, G. A., Rosa, G. J. M., Valente, B. D., Carvalheiro, R., Baldi, F., Garcia, D. A., et al. (2016). Genomic prediction of breeding values for carcass traits in Nellore cattle. *Genet. Sel. Evol.* 48, 7. doi:10.1186/s12711-016-0188-y.

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., et al. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Sci. Rep.* 9, 5364. doi:10.1038/s41598-019-41813-x.

Fonseca, P. A. de S., Id-Lahoucine, S., Reverter, A., Medrano, J. F., Fortes, M. S., Casellas, J., et al. (2018). Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS One* 13, e0205295. doi:10.1371/journal.pone.0205295.

Fontanesi, L. (2016). Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. *Anim. Front.* 6, 73. doi:10.2527/af.2016-0011.

Goldansaz, S. A., Markus, S., Berjanskii, M., Rout, M., Guo, A. C., Wang, Z., et al. (2020). Candidate serum metabolite biomarkers of residual feed intake and carcass merit in sheep. *J. Anim. Sci.* 98, 1–15. doi:10.1093/JAS/SKAA298.

Hanset, R., and Michaux, C. (1982). "Creatine and Creatinine Levels in Plasma, Red Cells and Muscles as Characteristics of Double-Muscled Cattle," in *Muscle Hypertrophy of Genetic Origin and its use to Improve Beef Production* (Springer Netherlands), 237–256. doi:10.1007/978-94-009-7550-7_21.

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., and Edwards, T. L. (2017). Population stratification in genetic association studies. *Curr. Protoc. Hum. Genet.* 95, 1.22.1-1.22.23. doi:10.1002/cphg.48.

Hwang, Y. H., and Joo, S. T. (2016). Fatty acid profiles of ten muscles from high and low marbled (Quality Grade 1++ and 2) hanwoo steers. *Korean J. Food Sci. Anim. Resour.* 36, 679–688. doi:10.5851/kosfa.2016.36.5.679.

Kazala, E. C., Lozeman, F. J., Mir, P. S., Aalhus, J. L., Schmutz, S. M., and Weselake, R. J. (2006). Fatty acid composition of muscle fat and enzymes of storage lipid synthesis in whole muscle from beef cattle. *Lipids* 41, 1049–1057. doi:10.1007/s11745-006-5055-0.

Kazala, E. C., Lozeman, F. J., Mir, P. S., Laroche, A., Bailey, D. R. C., and Weselake, R. J. (1999). Relationship of fatty acid composition to intramuscular fat content in beef from crossbred wagyu cattle. *J. Anim. Sci.* 77, 1717–1725. doi:10.2527/1999.7771717x.

Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., et al. (2020). Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle. *Front. Genet.* 11, 538600. doi:10.3389/fgene.2020.538600.

Li, J., Mukiibi, R., Wang, Y., Plastow, G. S., and Li, C. (2021). Identification of candidate genes and enriched biological functions for feed efficiency traits by integrating plasma metabolites and imputed whole genome sequence variants in beef cattle. *BMC Genomics*

22, 823. doi:10.1186/S12864-021-08064-5.

Lu, D., Sargolzaei, M., Kelly, M., Li, C., Voort, G. Vander, Wang, Z., et al. (2012). Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Genet.* 3, 152. doi:10.3389/fgene.2012.00152.

Lu, D., Sargolzaei, M., Kelly, M., Vander Voort, G., Wang, Z., Mandell, I., et al. (2013). Genome-wide association analyses for carcass quality in crossbred beef cattle. *BMC Genet.* 14, 80. doi:10.1186/1471-2156-14-80.

Mao, F., Chen, L., Vinsky, M., Okine, E., Wang, Z., Basarab, J., et al. (2013). Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci.* 91, 2067–2076. doi:10.2527/jas.2012-5470.

Matthews, J. O., Southern, L. L., Higbie, A. D., Persica, M. A., and Bidner, T. D. (2001). Effects of betaine on growth, carcass characteristics, pork quality, and plasma metabolites of finishing pigs. *J. Anim. Sci.* 79, 722–728. doi:10.2527/2001.793722x.

Medeiros de Oliveira Silva, R., Bonvino Stafuzza, N., de Oliveira Fragomeni, B., Miguel Ferreira de Camargo, G., Matos Ceacero, T., Noely dos Santos Gonçalves Cyrillo, J., et al. (2017). Genome-wide association study for carcass traits in an experimental Nelore cattle population. *PLoS One* 12, e0169860. doi:10.1371/journal.pone.0169860.

Mehrban, H., Lee, D. H., Moradi, M. H., Ilcho, C., Naserkheil, M., and Ibáñez-Escriche, N. (2017). Predictive performance of genomic selection methods for carcass traits in Hanwoo beef cattle: Impacts of the genetic architecture. *Genet. Sel. Evol.* 49, 1. doi:10.1186/s12711-016-0283-0.

Melzer, N., Wittenburg, D., and Repsilber, D. (2013). Integrating Milk Metabolite Profile Information for the Prediction of Traditional Milk Traits Based on SNP Information for Holstein Cows. *PLoS One* 8, e70256. doi:10.1371/journal.pone.0070256.

Meuwissen, T., Hayes, B., and Goddard, M. (2013). Accelerating improvement of livestock with genomic selection. *Annu. Rev. Anim. Biosci.* 1, 221–237. doi:10.1146/annurev-animal-031412-103705.

Meuwissen, T., van den Berg, I., and Goddard, M. (2021). On the use of whole-genome sequence data for across-breed genomic prediction and fine-scale mapping of QTL. *Genet. Sel. Evol.* 53, 19. doi:10.1186/s12711-021-00607-4.

Mierziak, J., Burgberger, M., and Wojtasik, W. (2021). 3-hydroxybutyrate as a metabolite and a signal molecule regulating processes of living organisms. *Biomolecules* 11, 402. doi:10.3390/biom11030402.

Mukiibi, R., Vinsky, M., Keogh, K. A., Stothard, P., Waters, S. M., and Li, C. (2018). Transcriptome analysis of liver tissues in Charolais steers with divergent marbling phenotypes. in *Proceedings of the World Congress on Genetics Applied to Livestock Production* (Auckland, New Zealand), 11.626.

Nkrumah, J. D., Keisler, D. H., Crews, D. H., Basarab, J. A., Wang, Z., Li, C., et al. (2007a). Genetic and phenotypic relationships of serum leptin concentration with performance, efficiency of gain, and carcass merit of feedlot cattle1. *J. Anim. Sci.* 85, 2147–2155.

doi:10.2527/jas.2006-764.

Nkrumah, J. D., Sherman, E. L., Li, C., Marques, E., Crews, D. H., Bartusiak, R., et al. (2007b). Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *J. Anim. Sci.* 85, 3170–3181. doi:10.2527/jas.2007-0234.

Ogawa, S., Matsuda, H., Taniguchi, Y., Watanabe, T., Kitamura, Y., Tabuchi, I., et al. (2017). Genomic prediction for carcass traits in Japanese Black cattle using single nucleotide polymorphism markers of different densities. *Anim. Prod. Sci.* 57, 1631. doi:10.1071/AN15696.

Patel, S. S., Molnar, M. Z., Tayek, J. A., Ix, J. H., Noori, N., Benner, D., et al. (2013). Serum creatinine as a marker of muscle mass in chronic kidney disease: Results of a cross-sectional study and review of literature. *J. Cachexia. Sarcopenia Muscle* 4, 19–29. doi:10.1007/s13539-012-0079-1.

Smith, Z. K., and Johnson, B. J. (2020). Mechanisms of steroidal implants to improve beef cattle growth: a review. *J. Appl. Anim. Res.* 48, 133–141. doi:10.1080/09712119.2020.1751642.

Snelling, W. M., Cushman, R. A., Keele, J. W., Maltecca, C., Thomas, M. G., Fortes, M. R. S., et al. (2013). BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection. *J. Anim. Sci.* 91, 537–552. doi:10.2527/JAS.2012-5784.

Todd, D. L., Roughsedge, T., and Woolliams, J. A. (2014). Genomic selection using beef commercial carcass phenotypes. *Animal* 8, 388–394. doi:10.1017/S1751731113002231.

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

Wang, X., Zhang, Y., Zhang, X., Wang, D., Jin, G., Li, B., et al. (2017). The comprehensive liver transcriptome of two cattle breeds with different intramuscular fat content. *Biochem. Biophys. Res. Commun.* 490, 1018–1025. doi:10.1016/j.bbrc.2017.06.157.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Weber, K. L., Welly, B. T., Van Eenennaam, A. L., Young, A. E., Porto-Neto, L. R., Reverter, A., et al. (2016). Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. *PLoS One* 11, e0152274. doi:10.1371/journal.pone.0152274.

Widmann, P., Reverter, A., Fortes, M. R. S., Weikard, R., Suhre, K., Hammon, H., et al. (2013). A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* 14, 798. doi:10.1186/1471-2164-14-798.

Wyss, M., and Kaddurah-Daouk, R. (2000). Creatine and creatinine metabolism. *Physiol. Rev.* 80, 1107–1213. doi:10.1152/physrev.2000.80.3.1107.

Xu, L., Gao, N., Wang, Z., Xu, L., Liu, Y., Chen, Y., et al. (2020). Incorporating genome annotation into genomic prediction for carcass traits in Chinese Simmental beef cattle.

Front. Genet. 11, 481. doi:10.3389/fgene.2020.00481.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi:10.1016/j.ajhg.2010.11.011.

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46, 100–106. doi:10.1038/ng.2876.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (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, 36. doi:10.1186/s12864-019-6362-1.

Zhang, H., Yin, L., Wang, M., Yuan, X., and Liu, X. (2019). Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Front. Genet.* 10, 189. doi:10.3389/fgene.2019.00189.

Table 4.1 A summary of metabolites associated with carcass merit traits in a multibreed population of beef cattle

Trait[1]	Metabolite[2]	*P*-value[3]	b[4]	V_m/V_P (%)[5]
HCW	3-hydroxybutyric acid	1.01E-02	-2.33E-01	1.92
	acetic acid	1.56E-02	-3.61E-02	1.69
	citric acid	2.35E-02	-1.10E-01	1.48
	choline	5.31E-02	4.03E-02	1.09
REA	creatinine	1.15E-02	2.50E-02	1.87
	L-glutamine	1.44E-02	5.51E-02	1.79
	succinic acid	1.90E-02	-3.46E-02	1.60
	pyruvic acid	5.32E-02	1.87E-02	1.56
	L-lactic acid	5.94E-02	4.61E-04	1.04
	3-hydroxybutyric acid	9.70E-02	-2.47E-02	0.80
AFAT	fumaric acid	4.05E-02	-7.90E-02	2.40
	methanol	4.81E-02	-5.15E-03	1.71
	D-glucose	5.01E-02	-8.64E-04	1.67
	glycerol	7.29E-02	9.33E-04	1.39
LMY	L-lactic acid	1.21E-02	2.73E-04	2.71
	creatinine	7.15E-02	7.41E-03	1.42
CMAR	succinic acid	1.53E-02	-2.18E-01	1.76
	fumaric acid	7.83E-02	-9.86E-01	1.36
	lysine	4.36E-02	1.94E-01	1.22
	glycine	4.55E-02	-4.71E-02	1.20
	choline	6.11E-02	-3.80E-02	1.05

[1]*HCW* hot carcass weight in kg, *REA* rib eye area in cm^2, *AFAT* average backfat thickness in mm, *LMY* lean meat yield in %, *CMAR* carcass marbling score from 100 (trace marbling) to 499 (more marbling)

[2]The unit of metabolite concentration is μM

[3]The significance level of regression analysis is *P*-value < 0.1

[4]*b* regression coefficient

[5]V_m/V_P: the proportion of phenotypic variance of carcass merit traits explained by associated metabolites (%)

Table 4.2 A summary of significant SNPs, the number of putative QTLs, and the number of candidate genes for metabolites associated with carcass merit traits in a multibreed population of beef cattle

Metabolite[1]	*P*-value range[2]	β range[3]	V_{SNP}/V_P range (%)[4]	V_{SNP}/V_P mean (%)[5]	No. of QTL[6]	No. of gene[7]
acetic acid	2.46E-12 – 9.97E-06	-259.54 – 181.91	4.01 – 10.95	5.05	31	49
citric acid	1.47E-06 – 9.75E-06	-29.80 – 37.62	3.57 – 4.95	4.05	15	15
choline	4.94E-07 – 9.90E-06	-89.13 – 84.25	3.85 – 5.43	4.61	13	23
D-glucose	6.82E-07 – 9.72E-06	-226.21 – 257.62	3.70 – 5.67	4.33	18	23
glycine	3.17E-06 – 9.54E-06	-68.80 – 75.32	3.97 – 4.65	4.31	9	10
glycerol	1.71E-07 – 9.76E-06	-457.22 – 355.60	4.05 – 6.67	4.93	21	29
fumaric acid	2.24E-07 – 9.83E-06	2.28 – 5.06	5.78 – 7.65	6.95	8	3
lysine	9.11E-09 – 9.80E-06	-17.77 – 20.56	3.88 – 7.13	4.82	15	20
L-lactic acid	2.24E-07 – 9.43E-06	-1076.62 – 1261.24	3.74 – 5.95	4.58	16	21
pyruvic acid	7.82E-08 – 9.99E-06	-42.51 – 47.85	6.03 – 9.88	6.90	18	32
succinic acid	3.32E-07 – 9.92E-06	-30.10 – 25.93	3.94 – 6.19	5.01	26	53
3-hydroxybutyric acid	8.55E-07 – 9.95E-06	11.24 – 19.78	4.00 – 5.13	4.46	12	19
creatinine	1.64E-07 – 9.86E-06	-26.59 – 31.04	3.78 – 6.31	4.67	17	22
L-glutamine	7.37E-07 – 9.90E-06	-11.76 – 11.53	4.06 – 5.30	4.66	13	13
methanol	3.15E-06 – 9.86E-06	-32.86 – 45.32	4.00 – 4.82	4.30	15	28

[1]The unit of metabolite concentration is μM
[2]The *P*-value range (minimum to maximum) of significant SNPs, the significance level is *P*-value < 1×10^{-5}
[3]*β range*: the range of allele substitution effect of each significant SNP
[4]*V_{SNP}/V_P range*: the range metabolite phenotypic variance explained by each significant SNP (%)
[5]*V_{SNP}/V_P mean*: the average of metabolite phenotypic variance explained by each significant SNP (%)
[6]*No. of QTL*: the number of putative QTLs identified for each metabolite
[7]*No. of gene*: the number of candidate gene identified for each metabolite

Table 4.3 Metabolites and their candidate genes associated with carcass merit traits in a multibreed population of beef cattle

Trait[1]	Metabolite[2]	Candidate gene
HCW	3-hydroxybutyric acid	*RNASE1, RNASE6, RNASE4, ANG2, CDH13, TRAF3, AMN, CDC42BPB, PGM2, SULT1E1, CSN1S1, CSN2, HSTN, FRAS1, ANXA3, LOX, SRFBP1, NTRK2, COL12A1*
	acetic acid	*PFN2, AGTR1, S100Z, CRHBP, AGGF1, LBH, YPEL5, ATAD2B, KLHL29, OR12K5, OR1B1, OR1L1, OR1L3, OR10W4, OR5B17, IZUMO2, MYH14, ZNF814, SNORA70, TMEM132E, HMGCLL1, GFRAL, TRAM2, TMEM14A, GSTA2, GPR139, UMOD, PDILT, ACSM5, ACSM2B, ACSM1, UQCRC2, PDZD9, MOSMO, VWA3A, IL21R, GTF3C1, KATNIP, ARMH3, HPS6, LDB1, PPRC1, SNORD22, DUSP5, SMC3, RBM20, KLRC1, APBB2, MAN2A1*
	citric acid	*SERPINE3, INTS6, ZNF667, ZNF583, USP32, CA4, ZNHIT3, MYO19, TRAF3, AMN, CDC42BPB, EDEM1, ARL8B, KLHL31, SLC28A3*
	choline	*HHAT, CDH8, PECAM1, MILR1, POLG2, DDX5, CEP95, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, LRP5, PPP6R3, CPT1A, MRPL21, IGHMBP2, MRGPRF, CACNG2, IFT27, PVALB, BICD1, PERP*
REA	creatinine	*ZBTB21, UMODL1, L3HYPDH, JKAMP, RTN1, PPP2R5E, MAML2, STOX2, ENPP6, IRF2, PRIMPOL, ACSL1, CENPU, RAB38, LUZP2, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, LEPROT, DNAJC6*
	L-glutamine	*MYO16, UBE2E2, DDX56, NPC1L1, NUDCD3, CAMK2B, TRIM24, SVOPL, ATP6V0A4, PPP3CC, SORBS3, PDLIM2, CCAR2*
	succinic acid	*GPR149, DHX36, ARHGEF26, PFN2, RNF13, HLTF, GYG1, AGTR1, ZIC1, ZIC4, SRSF5, SLC10A1, SMOC1, ACTR2, SPRED2, NDUFA8, MORN5, LHX6, RBM18, MRRF, PTGS1, OR1J2, OR1N2, OR1N1, OR1Q1, OR12K5, OR1B1, OR1L1, OR1L3, OR1AF3, OR1AF1, PDCL, RC3H2, ZBTB6, ZBTB26, RABGAP1, ZNF814, PLA2G2D1, PLA2G5, PLA2G2A, PLA2G2E, OTUD3, RNF186, TMCO4, LAP, DEFB13, RYR2, RBM47, NSUN7, APBB2, RAB28, NKX3-2, BOD1L1*
	pyruvic acid	*SLC49A4, SEMA5B, HMGB1, CDX2, PDX1, GSX1, CHST8, KCTD15, NTN1, GAS7, KCNJ2, KCNJ16, DNER, RAB3C, PRL, GALR1, MBP, ZNF236, ZNF516, CNDP2, DIPK1C, C24H18orf63, CYB5A, STAB2, NT5DC3, HSP90B1, C5H12orf73, TDG, GLT8D2, PHF21B, NUP50, RIMS1*
	L-lactic acid	*PLSCR1, AQP9, NEDD4, PRTG, PYGO1, CUX2, NOS1, FBXO21, SPPL3, HNF1A, C17H12orf43, OASL, FOXN4, ACACB, TMEM171, FCHO2, CD247, POU2F1, MACF1, NPFFR2, SGCD*
	3-hydroxybutyric acid	*RNASE1, RNASE6, RNASE4, ANG2, CDH13, TRAF3, AMN, CDC42BPB, PGM2, SULT1E1, CSN1S1, CSN2, HSTN, FRAS1, ANXA3, LOX, SRFBP1, NTRK2, COL12A1*
AFAT	fumaric acid	*CDH13, SLC17A6, PPP2R2B*
	methanol	*GRIK4, GRAMD1B, SCN3B, ZNF202, SMYD3, KIF26B, RGL1, CCDC92, DNAH10, PLA2G2A, PLA2G2E, OTUD3, MCTP2, XPC, TMEM43, CHCHD4, WNT7A, LUZP2, MARCHF1, PLAC8B, PLAC8A, COQ2, DENND4C, PLIN2, HAUS6, NEFM, NEFL, DOCK5*

	D-glucose	*TIAM1, KATNBL1, EMC7, CHRM5, AVEN, UBAC2, ST18, OR5M3, OR5M11, OR5AR1, INPP4B, PNKD, CATIP, SLC11A1, CTDSP1, VIL1, USP37, RARB, PGM5, TMEM252, WDR27, C9H6orf120, PHF10*
	glycerol	*DGKG, SNX31, ANKRD46, BCO2, PTS, C15H11orf34, WIPF1, CPEB4, C20H5orf47, NSG2, CCDC88C, PPP4R3A, RAB23, BAG2, ZNF451, KCNK17, KCNK16, KIF6, STOX2, THRB, CDK14, TRBV15, LOX, SRFBP1, FSTL4, JADE2, SAR1B, SEC24A, TJP2*
LMY	L-lactic acid	*PLSCR1, AQP9, NEDD4, PRTG, PYGO1, CUX2, NOS1, FBXO21, SPPL3, HNF1A, C17H12orf43, OASL, FOXN4, ACACB, TMEM171, FCHO2, CD247, POU2F1, MACF1, NPFFR2, SGCD*
	creatinine	*ZBTB21, UMODL1, L3HYPDH, JKAMP, RTN1, PPP2R5E, MAML2, STOX2, ENPP6, IRF2, PRIMPOL, ACSL1, CENPU, RAB38, LUZP2, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, LEPROT, DNAJC6*
CMAR	succinic acid	*GPR149, DHX36, ARHGEF26, PFN2, RNF13, HLTF, GYG1, AGTR1, ZIC1, ZIC4, SRSF5, SLC10A1, SMOC1, ACTR2, SPRED2, NDUFA8, MORN5, LHX6, RBM18, MRRF, PTGS1, OR1J2, OR1N2, OR1N1, OR1Q1, OR12K5, OR1B1, OR1L1, OR1L3, OR1AF3, OR1AF1, PDCL, RC3H2, ZBTB6, ZBTB26, RABGAP1, ZNF814, PLA2G2D1, PLA2G5, PLA2G2A, PLA2G2E, OTUD3, RNF186, TMCO4, LAP, DEFB13, RYR2, RBM47, NSUN7, APBB2, RAB28, NKX3-2, BOD1L1*
	fumaric acid	*CDH13, SLC17A6, PPP2R2B*
	lysine	*BTLA, ATG3, SLC35A5, CCDC80, CD200R1L, GTPBP8, NEPRO, BOC, SPICE1, SIDT1, FGF12, HS6ST3, FRMD5, MFHAS1, STYXL2, GPA33, DAB1, OR6C75, ITPR2, SSPN*
	glycine	*AQP9, PHLDB1, TREH, DDX6, EIF5, MARK3, SEM1, PINX1, SOX7, C8H8orf74*
	choline	*HHAT, CDH8, PECAM1, MILR1, POLG2, DDX5, CEP95, ALDH3B1, NDUFS8, TCIRG1, CHKA, KMT5B, LRP5, PPP6R3, CPT1A, MRPL21, IGHMBP2, MRGPRF, CACNG2, IFT27, PVALB, BICD1, PERP*

[1]*HCW* hot carcass weight in kg, *REA* rib eye area in cm^2, *AFAT* average backfat thickness in mm, *LMY* lean meat yield in %, *CMAR* carcass marbling score from 100 (trace marbling) to 499 (more marbling)
[2]The unit of metabolite concentration is μM

Table 4.4 Five topmost significantly enriched biological functions for carcass merit traits, and genes involved in functions

Trait[1]	Biological function	*P*-value range[2]	Genes involved in the biological function
HCW	Cell-To-Cell Signaling and Interaction	8.91E-04 – 2.52E-02	*AGTR1, CACNG2, CDH13, CDH8, KLRC1, LOX, MAN2A1, NTRK2, PECAM1, PERP, PVALB, TRAF3, UMOD*
	Molecular Transport	8.91E-04 – 2.52E-02	*AGTR1, CHKA, CPT1A, DDX5, HPS6, IGHMBP2, IL21R, KLRC1, LRP5, NTRK2, PECAM1, PVALB, SLC28A3, SULT1E1, TRAF3, UMOD*
	Small Molecule Biochemistry	8.91E-04 – 2.52E-02	*AGTR1, CHKA, CPT1A, DDX5, HMGCLL1, IGHMBP2, IL21R, KLRC1, LOX, LRP5, MAN2A1, NTRK2, PECAM1, PFN2, PGM2, PVALB, SLC28A3, SULT1E1, TRAF3, UMOD*
	Lipid Metabolism	9.14E-04 – 2.52E-02	*AGTR1, CHKA, CPT1A, DDX5, HMGCLL1, IGHMBP2, IL21R, LRP5, NTRK2, PVALB, SULT1E1, TRAF3*
	Carbohydrate Metabolism	1.35E-03 – 2.52E-02	*AGTR1, CPT1A, LRP5, MAN2A1, PGM2*
REA	Molecular Transport	1.03E-06 – 6.53E-03	*ACACB, ACSL1, AGTR1, AMN, AQP9, ATP6V0A4, CAMK2B, CDX2, CHKA, CHST8, CSN2, DDX56, FRAS1, GALR1, GAS7, GSX1, HMGB1, HNF1A, IRF2, KCNJ16, KCNJ2, MBP, NEDD4, NOS1, NPC1L1, NTN1, NTRK2, NUP50, PFN2, PLA2G2A, PLA2G2E, PLA2G5, PLSCR1, PPP3CC, PRL, PTGS1, RAB3C, RIMS1, RYR2, SLC10A1, SORBS3, SPRED2, SRSF5, SULT1E1, TCIRG1, TRAF3, ZBTB21, ZIC1*
	Lipid Metabolism	3.27E-05 – 6.53E-03	*ACACB, ACSL1, AGTR1, ALDH3B1, AQP9, CHKA, CHST8, CYB5A, ENPP6, GALR1, GAS7, HMGB1, HNF1A, IRF2, MBP, NOS1, NPC1L1, NTN1, NTRK2, PLA2G2A, PLA2G2E, PLA2G5, PLSCR1, PRL, PTGS1, RIMS1, SLC10A1, SPRED2, SULT1E1, TRAF3, ZBTB21*
	Small Molecule Biochemistry	3.27E-05 – 6.53E-03	*ACACB, ACSL1, AGTR1, ALDH3B1, AQP9, CHKA, CHST8, CNDP2, CYB5A, ENPP6, GALR1, GAS7, GSX1, HMGB1, HNF1A, IRF2, LOX, MBP, NOS1, NPC1L1, NTN1, NTRK2, PFN2, PLA2G2A, PLA2G2E, PLA2G5, PLSCR1, PPP3CC, PRL, PTGS1, RIMS1, RYR2, SLC10A1, SORBS3, SPRED2, STAB2, SULT1E1, TDG, TRAF3, ZBTB21*
	Cellular Assembly and Organization	9.97E-05 – 6.53E-03	*ACTR2, CAMK2B, CDC42BPB, CDH13, CUX2, DNAJC6, DNER, FCHO2, GAS7, HMGB1, KCNJ2, LOX, MBP, MYO16, NOS1, NTN1, NTRK2, NUDCD3, PFN2, PLSCR1, PPP3CC, PRL, RAB28, RAB38, RIMS1, RYR2*
	Cell-To-Cell Signaling and Interaction	1.5E-04 – 6.53E-03	*AGTR1, CAMK2B, CD247, CDH13, CUX2, GALR1, HMGB1, HSP90B1, NOS1, NTN1, NTRK2, PFN2, PLA2G5, PPP3CC, PRL, RIMS1, SORBS3, TRAF3*
AFAT	Cell-To-Cell Signaling and Interaction	1.63E-05 – 2.67E-02	*CDH13, CHRM5, GRIK4, LOX, MARCHF1, NEFL, NEFM, PNKD, PTS, RARB, SLC17A6, THRB, TIAM1, TJP2, WNT7A, XPC*

	Drug Metabolism	1.63E-05 – 1.34E-02	*CHRM5, PNKD, PTS, SLC17A6*
	Molecular Transport	1.63E-05 – 2.67E-02	*CDK14, CHCHD4, CHRM5, GRIK4, INPP4B, KCNK16, KCNK17, NEFM, PLA2G2A, PLA2G2E, PLIN2, PNKD, PPP2R2B, PTS, SCN3B, SEC24A, SLC11A1, SLC17A6, TJP2, WNT7A, XPC, ZNF202*
	Small Molecule Biochemistry	1.63E-05 – 2.34E-02	*BCO2, CHRM5, COQ2, DGKG, GRIK4, INPP4B, LOX, PLA2G2A, PLA2G2E, PLIN2, PNKD, PTS, SLC11A1, SLC17A6, THRB, WNT7A, XPC*
	Cellular Assembly and Organization	3.36E-05 – 2.67E-02	*CATIP, CDH13, CHCHD4, CHRM5, CPEB4, DGKG, DOCK5, KATNBL1, LOX, NEFL, NEFM, RAB23, RARB, TIAM1, TJP2, VIL1, WIPF1, WNT7A, XPC*
LMY	Lipid Metabolism	1.18E-04 – 3.44E-02	*ACACB, ACSL1, ALDH3B1, AQP9, CHKA, ENPP6, HNF1A, NOS1, PLSCR1*
	Molecular Transport	1.18E-04 – 3.8E-02	*ACACB, ACSL1, AQP9, CHKA, HNF1A, NEDD4, NOS1, NPFFR2, PLSCR1, TCIRG1, ZBTB21*
	Nucleic Acid Metabolism	1.18E-04 – 2.91E-02	*ACACB, ACSL1, AQP9, NOS1*
	Small Molecule Biochemistry	1.18E-04 – 3.44E-02	*ACACB, ACSL1, ALDH3B1, AQP9, CHKA, ENPP6, HNF1A, IRF2, NOS1, NPFFR2, PLSCR1, RAB38*
	Amino Acid Metabolism	1.84E-03 – 1.64E-02	*HNF1A, NOS1, RAB38*
CMAR	Lipid Metabolism	1.57E-05 – 2.22E-02	*AGTR1, AQP9, CCDC80, CHKA, CPT1A, DAB1, DDX5, IGHMBP2, LRP5, PLA2G2A, PLA2G2E, PLA2G5, PTGS1, PVALB, SLC10A1, SPRED2, SSPN*
	Molecular Transport	1.57E-05 – 2.36E-02	*AGTR1, AQP9, CCDC80, CHKA, CPT1A, DAB1, DDX5, IGHMBP2, LRP5, PECAM1, PLA2G2A, PLA2G2E, PLA2G5, PTGS1, PVALB, SLC10A1, SLC17A6, SPRED2*
	Small Molecule Biochemistry	1.57E-05 – 2.22E-02	*AGTR1, AQP9, CCDC80, CHKA, CPT1A, DAB1, DDX5, IGHMBP2, LRP5, PECAM1, PFN2, PLA2G2A, PLA2G2E, PLA2G5, PTGS1, PVALB, SLC10A1, SLC17A6, SPRED2, SSPN, TREH*
	Carbohydrate Metabolism	4.1E-04 – 1.34E-02	*AGTR1, AQP9, CHKA, CPT1A, GYG1, LRP5, NKX3-2, PDCL, PLA2G2A, PLA2G2E, PLA2G5, TREH*
	Cell-To-Cell Signaling and Interaction	2.47E-03 – 1.78E-02	*AGTR1, CACNG2, CDH13, CDH8, PECAM1, PERP, PFN2, PLA2G5, PVALB, SLC17A6*

[1]*HCW* hot carcass weight in kg, *REA* rib eye area in cm^2, *AFAT* average backfat thickness in mm, *LMY* lean meat yield in %, *CMAR* carcass marbling score from 100 (trace marbling) to 499 (more marbling)

[2]The *P*-value range (minimum to maximum) of significant biological functions, the significance level is *P*-value < 0.05

Table 4.S1 Uniquely common candidate genes for carcass merit traits in a multibreed population of beef cattle

Trait	Number of common genes	Common gene name
AFAT, CMAR, HCW, REA	1	*CDH13*
CMAR, HCW, LMY, REA	5	*KMT5B, NDUFS8, ALDH3B1, CHKA, TCIRG1*
CMAR, HCW, REA	8	*PFN2, APBB2, OR1L1, ZNF814, OR12K5, AGTR1, OR1L3, OR1B1*
AFAT, HCW, REA	2	*SRFBP1, LOX*
AFAT, CMAR, REA	3	*PLA2G2A, OTUD3, PLA2G2E*
AFAT, LMY, REA	2	*LUZP2, STOX2*
CMAR, LMY, REA	1	*AQP9*
HCW, REA	16	*PGM2, RNASE1, SULT1E1, FRAS1, CSN1S1, ANG2, AMN, RNASE6, NTRK2, CDC42BPB, ANXA3, HSTN, CSN2, TRAF3, RNASE4, COL12A1*
CMAR, HCW	18	*MILR1, IGHMBP2, PPP6R3, MRPL21, DDX5, PECAM1, POLG2, BICD1, LRP5, CACNG2, CPT1A, PVALB, MRGPRF, CDH8, CEP95, HHAT, PERP, IFT27*
LMY, REA	35	*DNAJC6, L3HYPDH, TMEM171, RAB38, PRTG, MACF1, LEPROT, NPFFR2, FBXO21, HNF1A, RTN1, PRIMPOL, ENPP6, ACACB, NOS1, MAML2, JKAMP, C17H12orf43, NEDD4, UMODL1, PYGO1, PPP2R5E, SPPL3, ZBTB21, FOXN4, POU2F1, PLSCR1, ACSL1, CUX2, OASL, FCHO2, CENPU, CD247, IRF2, SGCD*
CMAR, REA	42	*RBM18, RNF186, MRRF, OR1N1, ZIC1, HLTF, ARHGEF26, SLC10A1, OR1J2, PDCL, PLA2G5, RABGAP1, ZBTB26, OR1AF1, NSUN7, RAB28, NDUFA8, PTGS1, TMCO4, GYG1, BOD1L1, RNF13, SPRED2, SRSF5, GPR149, OR1Q1, PLA2G2D1, NKX3-2, RBM47, LHX6, DEFB13, OR1N2, MORN5, DHX36, SMOC1, ZBTB6, RYR2, OR1AF3, LAP, ZIC4, ACTR2, RC3H2*
AFAT, CMAR	2	*PPP2R2B, SLC17A6*

Table 4.S2 Enriched biological functions significantly associated with HCW in a multibreed population of beef cattle

Biological function	*P*-value range	Genes involved in the biological function
Cell-To-Cell Signaling and Interaction	8.91E-04-2.52E-02	*AGTR1,CACNG2,CDH13,CDH8,KLRC1,LOX,MAN2A1,NTRK2,PECAM1,PERP, PVALB,TRAF3,UMOD*
Molecular Transport	8.91E-04-2.52E-02	*AGTR1,CHKA,CPT1A,DDX5,HPS6,IGHMBP2,IL21R,KLRC1,LRP5,NTRK2, PECAM1,PVALB,SLC28A3,SULT1E1,TRAF3,UMOD*
Small Molecule Biochemistry	8.91E-04-2.52E-02	*AGTR1,CHKA,CPT1A,DDX5,HMGCLL1,IGHMBP2,IL21R,KLRC1,LOX,LRP5, MAN2A1,NTRK2,PECAM1,PFN2,PGM2,PVALB,SLC28A3,SULT1E1,TRAF3, UMOD*
Lipid Metabolism	9.14E-04-2.52E-02	*AGTR1,CHKA,CPT1A,DDX5,HMGCLL1,IGHMBP2,IL21R,LRP5,NTRK2,PVALB, SULT1E1,TRAF3*
Carbohydrate Metabolism	1.35E-03-2.52E-02	*AGTR1,CPT1A,LRP5,MAN2A1,PGM2*
Cell Morphology	1.53E-03-2.52E-02	*AGTR1,CDH13,MYH14,NTRK2,PECAM1,PERP,PFN2,POLG2,PVALB,SULT1E1*
Cellular Function and Maintenance	1.53E-03-2.11E-02	*ARL8B,CACNG2,CDH13,CDH8,COL12A1,HPS6,IL21R,LRP5,MAN2A1,MYH14, NTRK2,PECAM1,PFN2,POLG2,PPRC1,PVALB,TCIRG1,TRAF3*
Cell Death and Survival	2.62E-03-2.11E-02	*AGTR1,ALDH3B1,ARL8B,CA4,KLRC1,LRP5,NTRK2,TRAF3,UMOD*
Cellular Compromise	2.62E-03-1.99E-02	*AGTR1,ALDH3B1,ANXA3,ARL8B,CA4,KLRC1,NTRK2,PECAM1,PGM2,TCIRG1, YPEL5*
Amino Acid Metabolism	4.25E-03-8.48E-03	*LOX*
Cell Cycle	4.25E-03-2.11E-02	*CDH13,CHKA,LBH,MYH14,SMC3*
Cellular Assembly and Organization	4.25E-03-2.52E-02	*ARL8B,CACNG2,CDH13,CDH8,HPS6,IFT27,KMT5B,LOX,MAN2A1,MYH14, NTRK2,PECAM1,PERP,PFN2,POLG2,PPRC1,PVALB,SMC3,UMOD*
Cellular Development	4.25E-03-2.52E-02	*AGTR1,CACNG2,CDH13,CHKA,DUSP5,IL21R,LBH,LOX,NTRK2,PECAM1, PFN2,SMC3,TRAF3*
Cellular Growth and Proliferation	4.25E-03-2.52E-02	*AGTR1,CACNG2,CDH13,DUSP5,IL21R,LDB1,LOX,NTRK2,PECAM1,PFN2, SMC3,TRAF3*
Cellular Movement	4.25E-03-2.52E-02	*AGTR1,ANXA3,CA4,CDH13,LOX,MAN2A1,MYH14,NTRK2,PECAM1,SULT1E1, UMOD*
Drug Metabolism	4.25E-03-1.22E-02	*ACSM1,ACSM2B,GSTA2,PFN2,PVALB,SULT1E1*
Nucleic Acid Metabolism	4.25E-03-1.69E-02	*CPT1A,HMGCLL1,PFN2,SLC28A3,UMOD*
Post-Translational Modification	4.25E-03-8.48E-03	*LOX*
RNA Post-Transcriptional Modification	4.25E-03-4.25E-03	*DDX5*
Energy Production	8.48E-03-2.52E-02	*CPT1A,LOX*
Gene Expression	8.48E-03-1.78E-02	*DDX5,GTF3C1,IGHMBP2,SMC3*
Cell Signaling	1.02E-02-2.11E-02	*AGTR1,CDH13,PECAM1,PVALB*

Free Radical Scavenging	1.27E-02-1.27E-02	*AGTR1*
Protein Synthesis	1.27E-02-1.69E-02	*PECAM1,PFN2*
Vitamin and Mineral Metabolism	1.27E-02-2.11E-02	*PVALB,SULT1E1*
Protein Trafficking	1.69E-02-1.69E-02	*TRAF3*

Table 4.S3 Enriched biological functions significantly associated with REA in a multibreed population of beef cattle

Biological function	*P*-value range	Genes involved in the biological function
Molecular Transport	1.03E-06-6.53E-03	*ACACB,ACSL1,AGTR1,AMN,AQP9,ATP6V0A4,CAMK2B,CDX2,CHKA,CHST8,CSN2, DDX56,FRAS1,GALR1,GAS7,GSX1,HMGB1,HNF1A,IRF2,KCNJ16,KCNJ2,MBP,NEDD4, NOS1,NPC1L1,NTN1,NTRK2,NUP50,PFN2,PLA2G2A,PLA2G2E,PLA2G5,PLSCR1, PPP3CC,PRL,PTGS1,RAB3C,RIMS1,RYR2,SLC10A1,SORBS3,SPRED2,SRSF5,SULT1E1, TCIRG1,TRAF3,ZBTB21,ZIC1*
Lipid Metabolism	3.27E-05-6.53E-03	*ACACB,ACSL1,AGTR1,ALDH3B1,AQP9,CHKA,CHST8,CYB5A,ENPP6,GALR1,GAS7, HMGB1,HNF1A,IRF2,MBP,NOS1,NPC1L1,NTN1,NTRK2,PLA2G2A,PLA2G2E,PLA2G5, PLSCR1,PRL,PTGS1,RIMS1,SLC10A1,SPRED2,SULT1E1,TRAF3,ZBTB21*
Small Molecule Biochemistry	3.27E-05-6.53E-03	*ACACB,ACSL1,AGTR1,ALDH3B1,AQP9,CHKA,CHST8,CNDP2,CYB5A,ENPP6,GALR1, GAS7,GSX1,HMGB1,HNF1A,IRF2,LOX,MBP,NOS1,NPC1L1,NTN1,NTRK2,PFN2, PLA2G2A,PLA2G2E,PLA2G5,PLSCR1,PPP3CC,PRL,PTGS1,RIMS1,RYR2,SLC10A1, SORBS3,SPRED2,STAB2,SULT1E1,TDG,TRAF3,ZBTB21*
Cellular Assembly and Organization	9.97E-05-6.53E-03	*ACTR2,CAMK2B,CDC42BPB,CDH13,CUX2,DNAJC6,DNER,FCHO2,GAS7,HMGB1, KCNJ2,LOX,MBP,MYO16,NOS1,NTN1,NTRK2,NUDCD3,PFN2,PLSCR1,PPP3CC,PRL, RAB28,RAB38,RIMS1,RYR2*
Cell-To-Cell Signaling and Interaction	1.5E-04-6.53E-03	*AGTR1,CAMK2B,CD247,CDH13,CUX2,GALR1,HMGB1,HSP90B1,NOS1,NTN1,NTRK2, PFN2,PLA2G5,PPP3CC,PRL,RIMS1,SORBS3,TRAF3*
Carbohydrate Metabolism	2.92E-04-6.53E-03	*ACACB,AGTR1,AQP9,CHKA,CHST8,CYB5A,ENPP6,GALR1,GYG1,HMGB1,HNF1A,NKX3-2,PDCL,PDX1,PGM2,PLA2G2A,PLA2G2E,PLA2G5,PLSCR1,PRL,RYR2,STAB2*
Energy Production	3.71E-04-3.71E-04	*ACACB,HNF1A,PRL,RYR2*
Cell Death and Survival	3.84E-04-5.32E-03	*CAMK2B,GAS7,HMGB1,MBP,NTRK2,PRL,RNF13,RTN1*
Cellular Compromise	4.18E-04-6.53E-03	*IRF2,MBP,NTN1*
Cellular Development	5.65E-04-6.53E-03	*AGTR1,CAMK2B,CUX2,DNER,GAS7,HMGB1,HNF1A,HSP90B1,KCNJ2,LOX,MBP,MYO16, NOS1,NTN1,NTRK2,PDX1,PFN2,PLSCR1,RIMS1,RNASE1,RYR2,TRAF3,ZIC1,ZIC4*
Gene Expression	6.22E-04-3.63E-03	*ACACB,ACTR2,APBB2,CCAR2,CDH13,CDX2,CSN2,CUX2,DHX36,FOXN4,GALR1,GSX1, HLTF,HMGB1,HNF1A,IRF2,LOX,MAML2,MBP,NKX3-2,NOS1,NTRK2,PDX1,PLSCR1,*

		POU2F1,PRL,PYGO1,RC3H2,RIMS1,SORBS3,TCIRG1, TDG,TRAF3,TRIM24,UMODL1,ZBTB21,ZIC1,ZNF516
Cell Morphology	6.25E-04-6.53E-03	*ACTR2,CAMK2B,CDC42BPB,CDH13,CDX2,CUX2,DNER,FOXN4,GAS7,HMGB1,KCNJ2, LOX,MBP,MYO16,NOS1,NTN1,NTRK2,NUDCD3,PFN2,PRL,RIMS1,RYR2,ZIC4*
Protein Synthesis	1.07E-03-3.66E-03	*CHST8,NOS1,NTRK2,PRL*
Cellular Function and Maintenance	1.47E-03-6.53E-03	*ACACB,ACSL1,ACTR2,AGTR1,AMN,ATP6V0A4,CAMK2B,CCAR2,CD247,CDC42BPB, CDH13,CUX2,DNAJC6,DNER,FCHO2,GAS7,HMGB1,HNF1A,HSP90B1,IRF2,KCNJ2,LOX, MBP,MYO16,NEDD4,NOS1,NTN1,NTRK2,NUDCD3,PDX1,PFN2,PLA2G2A,PLA2G5, PLSCR1,PPP3CC,PRL,PTGS1,RC3H2,RIMS1,RNASE1,RTN1,RYR2,SGCD,SPPL3,SPRED2, STAB2,TCIRG1,TRAF3,TRIM24*
Cellular Growth and Proliferation	1.47E-03-6.53E-03	*AGTR1,CAMK2B,CCAR2,CUX2,DNER,GAS7,HLTF,HMGB1,HNF1A,IRF2,KCNJ2,LOX, MBP,MYO16,NOS1,NTN1,NTRK2,PDX1,PFN2,PLSCR1,PRL,RIMS1,RYR2,TRAF3,ZIC1*
Nucleic Acid Metabolism	1.48E-03-6.53E-03	*ACACB,ACSL1,PFN2,TDG*
Cellular Movement	2.89E-03-6.53E-03	*CDH13,FCHO2,HMGB1,LOX,NTN1,NTRK2,PDCL,PFN2,PPP3CC,PRL,RIMS1,SULT1E1*
Cell Signaling	3.66E-03-6.53E-03	*IRF2,NOS1,RYR2*
Vitamin and Mineral Metabolism	3.66E-03-3.66E-03	*NOS1,RYR2*
Drug Metabolism	5.59E-03-6.53E-03	*HMGB1,IRF2,PFN2,PTGS1,SULT1E1*
Amino Acid Metabolism	6.53E-03-6.53E-03	*CNDP2,LOX,NOS1,PRL*
Cell Cycle	6.53E-03-6.53E-03	*CDH13,CHKA,HMGB1*
DNA Replication, Recombination, and Repair	6.53E-03-6.53E-03	*IRF2*
Post-Translational Modification	6.53E-03-6.53E-03	*LOX*

Table 4.S4 Enriched biological functions significantly associated with AFAT in a multibreed population of beef cattle

Biological function	*P*-value range	Genes involved in the biological function
Cell-To-Cell Signaling and Interaction	1.63E-05-2.67E-02	*CDH13,CHRM5,GRIK4,LOX,MARCHF1,NEFL,NEFM,PNKD,PTS,RARB, SLC17A6,THRB,TIAM1,TJP2,WNT7A,XPC*
Drug Metabolism	1.63E-05-1.34E-02	*CHRM5,PNKD,PTS,SLC17A6*
Molecular Transport	1.63E-05-2.67E-02	*CDK14,CHCHD4,CHRM5,GRIK4,INPP4B,KCNK16,KCNK17,NEFM,PLA2G2A, PLA2G2E,PLIN2,PNKD,PPP2R2B,PTS,SCN3B,SEC24A,SLC11A1,SLC17A6, TJP2,WNT7A,XPC,ZNF202*
Small Molecule Biochemistry	1.63E-05-2.34E-02	*BCO2,CHRM5,COQ2,DGKG,GRIK4,INPP4B,LOX,PLA2G2A,PLA2G2E,PLIN2, PNKD,PTS,SLC11A1,SLC17A6,THRB,WNT7A,XPC*
Cellular Assembly and Organization	3.36E-05-2.67E-02	*CATIP,CDH13,CHCHD4,CHRM5,CPEB4,DGKG,DOCK5,KATNBL1,LOX,NEFL, NEFM,RAB23,RARB,TIAM1,TJP2,VIL1,WIPF1,WNT7A,XPC*
Cellular Development	6.72E-05-2.67E-02	*CDH13,CHRM5,CTDSP1,DOCK5,INPP4B,LOX,LUZP2,NEFL,NEFM,RAB23, RARB,SLC17A6,THRB,TIAM1,WNT7A*
Cellular Function and Maintenance	6.72E-05-2.67E-02	*CDH13,CHRM5,CPEB4,DGKG,DOCK5,KATNBL1,LOX,MARCHF1,NEFL, NEFM,RAB23,THRB,TIAM1,VIL1,WIPF1,WNT7A,XPC*
Cellular Growth and Proliferation	6.72E-05-2.67E-02	*CDH13,CHRM5,CTDSP1,DOCK5,INPP4B,LOX,LUZP2,NEFL,NEFM,RARB, SLC11A1,SLC17A6,TIAM1,WNT7A*
Cell Morphology	1.67E-04-2.67E-02	*CCDC92,CDH13,CHRM5,CPEB4,DGKG,DOCK5,KIF26B,LOX,NEFL,NEFM, PLA2G2A,PLIN2,RAB23,RARB,THRB,TIAM1,TJP2,WIPF1,WNT7A,ZNF451*
Amino Acid Metabolism	3.37E-03-2.34E-02	*LOX,SLC11A1,SLC17A6*
Antigen Presentation	3.37E-03-3.37E-03	*MARCHF1*
Carbohydrate Metabolism	3.37E-03-2.34E-02	*CHRM5,COQ2,DGKG,INPP4B,PLA2G2A,PLA2G2E*
Cell Cycle	3.37E-03-2.01E-02	*CDH13,CDK14,LOX,NEFL,RARB,SLC11A1,THRB,XPC*
Cellular Movement	3.37E-03-2.51E-02	*CDH13,DOCK5,INPP4B,KIF26B,LOX,MARCHF1,THRB,TIAM1*
Lipid Metabolism	3.37E-03-2.34E-02	*BCO2,CHRM5,DGKG,INPP4B,PLA2G2A,PLA2G2E,PLIN2,THRB,XPC*
Post-Translational Modification	3.37E-03-2.67E-02	*CHCHD4,LOX,PLIN2*
Protein Trafficking	3.37E-03-3.37E-03	*PPP2R2B*
Vitamin and Mineral Metabolism	3.37E-03-1.01E-02	*BCO2,PLIN2*
Cell Death and Survival	6.74E-03-2.01E-02	*NEFL,NEFM,PLA2G2A,PLIN2,SLC11A1,THRB,WIPF1,XPC*
Energy Production	6.74E-03-6.74E-03	*LOX*
Cell Signaling	7.3E-03-7.3E-03	*CDH13,TIAM1*
Cellular Compromise	1.01E-02-1.68E-02	*DOCK5,NEFL,NEFM,PLA2G2A,SLC11A1,WIPF1*
DNA Replication, Recombination, and Repair	1.01E-02-1.68E-02	*CHCHD4,XPC*
Gene Expression	1.01E-02-1.19E-02	*LOX,RARB,THRB*
Nucleic Acid Metabolism	1.34E-02-1.34E-02	*XPC*
Protein Folding	2.67E-02-2.67E-02	*CHCHD4*

Table 4.S5 Enriched biological functions significantly associated with LMY in a multibreed population of beef cattle

Biological function	*P*-value range	Genes involved in the biological function
Lipid Metabolism	1.18E-04-3.44E-02	*ACACB,ACSL1,ALDH3B1,AQP9,CHKA,ENPP6,HNF1A,NOS1,PLSCR1*
Molecular Transport	1.18E-04-3.8E-02	*ACACB,ACSL1,AQP9,CHKA,HNF1A,NEDD4,NOS1,NPFFR2,PLSCR1, TCIRG1,ZBTB21*
Nucleic Acid Metabolism	1.18E-04-2.91E-02	*ACACB,ACSL1,AQP9,NOS1*
Small Molecule Biochemistry	1.18E-04-3.44E-02	*ACACB,ACSL1,ALDH3B1,AQP9,CHKA,ENPP6,HNF1A,IRF2,NOS1, NPFFR2,PLSCR1,RAB38*
Amino Acid Metabolism	1.84E-03-1.64E-02	*HNF1A,NOS1,RAB38*
Carbohydrate Metabolism	1.84E-03-3.44E-02	*ACACB,ACSL1,AQP9,CHKA,ENPP6,HNF1A,PLSCR1*
Cell Cycle	1.84E-03-3.75E-02	*CD247,CHKA,HNF1A,IRF2,PLSCR1,POU2F1,TCIRG1*
Cell Morphology	1.84E-03-3.44E-02	*AQP9,CD247,FOXN4,HNF1A,IRF2,NEDD4,NOS1,PLSCR1,SGCD*
Cell Signaling	1.84E-03-2.55E-02	*IRF2,NEDD4,NOS1,OASL,PLSCR1*
Cellular Assembly and Organization	1.84E-03-3.26E-02	*ACACB,CD247,DNAJC6,FCHO2,HNF1A,KMT5B,NEDD4,NOS1, PLSCR1,POU2F1,RAB38,RTN1,SGCD*
Cellular Compromise	1.84E-03-3.26E-02	*ACSL1,CD247,HNF1A,IRF2,NOS1,PLSCR1,PRTG,RTN1,TCIRG1*
Cellular Development	1.84E-03-3.78E-02	*CD247,CHKA,FOXN4,HNF1A,IRF2,NOS1,PLSCR1,PRTG,TCIRG1*
Cellular Function and Maintenance	1.84E-03-3.62E-02	*ACACB,ACSL1,AQP9,CD247,FCHO2,HNF1A,IRF2,NEDD4,NOS1, PLSCR1,RTN1,SGCD,SPPL3,TCIRG1*
Cellular Growth and Proliferation	1.84E-03-3.78E-02	*CD247,HNF1A,IRF2,NOS1,PLSCR1,POU2F1,TCIRG1*
DNA Replication, Recombination, and Repair	1.84E-03-3.44E-02	*HNF1A,IRF2,NOS1,POU2F1,PRIMPOL,RTN1*
Cell Death and Survival	3.68E-03-3.8E-02	*CD247,HNF1A,IRF2,NOS1,PLSCR1,POU2F1,SGCD*
Cell-To-Cell Signaling and Interaction	3.68E-03-3.44E-02	*CD247,IRF2,NOS1,NPFFR2,RTN1*
Cellular Movement	3.68E-03-1.28E-02	*HNF1A,MACF1,NEDD4,NOS1*
Energy Production	3.68E-03-2.91E-02	*ACACB,ACSL1,HNF1A,NOS1*
Post-Translational Modification	3.68E-03-2.37E-02	*HNF1A,NOS1*
Gene Expression	4.94E-03-3.44E-02	*ACACB,CUX2,FOXN4,HNF1A,IRF2,MAML2,NOS1,PLSCR1,POU2F1, PYGO1,TCIRG1,UMODL1,ZBTB21*
Vitamin and Mineral Metabolism	7.34E-03-2.55E-02	*ACACB,HNF1A,NOS1*
Protein Synthesis	8.02E-03-3.44E-02	*ACACB,HNF1A,IRF2,NOS1,NPFFR2,POU2F1,SGCD,TCIRG1*
Drug Metabolism	2.37E-02-2.37E-02	*ACSL1*

Table 4.S6 Enriched biological functions significantly associated with CMAR in a multibreed population of beef cattle

Biological function	*P*-value range	Genes involved in the biological function
Lipid Metabolism	1.57E-05-2.22E-02	*AGTR1,AQP9,CCDC80,CHKA,CPT1A,DAB1,DDX5,IGHMBP2,LRP5, PLA2G2A,PLA2G2E,PLA2G5,PTGS1,PVALB,SLC10A1,SPRED2,SSPN*
Molecular Transport	1.57E-05-2.36E-02	*AGTR1,AQP9,CCDC80,CHKA,CPT1A,DAB1,DDX5,IGHMBP2,LRP5, PECAM1,PLA2G2A,PLA2G2E,PLA2G5,PTGS1,PVALB,SLC10A1,SLC17A6, SPRED2*
Small Molecule Biochemistry	1.57E-05-2.22E-02	*AGTR1,AQP9,CCDC80,CHKA,CPT1A,DAB1,DDX5,IGHMBP2,LRP5, PECAM1,PFN2,PLA2G2A,PLA2G2E,PLA2G5,PTGS1,PVALB,SLC10A1, SLC17A6,SPRED2,SSPN,TREH*
Carbohydrate Metabolism	4.1E-04-1.34E-02	*AGTR1,AQP9,CHKA,CPT1A,GYG1,LRP5,NKX3-2,PDCL,PLA2G2A, PLA2G2E,PLA2G5,TREH*
Cell-To-Cell Signaling and Interaction	2.47E-03-1.78E-02	*AGTR1,CACNG2,CDH13,CDH8,PECAM1,PERP,PFN2,PLA2G5,PVALB, SLC17A6*
Cell Death and Survival	2.84E-03-1.73E-02	*FGF12,PLA2G2A,PTGS1,RNF186*
Cell Signaling	3.81E-03-2.22E-02	*AGTR1,BOC,CDH13,HHAT,IFT27,ITPR2,NDUFA8,NDUFS8,PECAM1, PTGS1,PVALB,ZIC1*
Cell Cycle	3.93E-03-2.22E-02	*BTLA,CDH13,CHKA,ZIC1*
Cell Morphology	4.47E-03-1.95E-02	*AGTR1,ARHGEF26,ATG3,BOC,CDH13,HHAT,LHX6,PECAM1,PERP,PFN2, PLA2G2A,POLG2,PVALB,RYR2,ZIC4*
Cellular Assembly and Organization	4.47E-03-2.22E-02	*ACTR2,APBB2,ATG3,BOC,CDH13,CDH8,DAB1,DDX6,HLTF,IFT27,ITPR2, KMT5B,PERP,PFN2,POLG2,PVALB,RAB28,RYR2,SLC17A6*
Cellular Development	4.47E-03-2.22E-02	*AGTR1,ARHGEF26,BOC,BTLA,CDH13,CHKA,LHX6,PECAM1,PFN2, SLC17A6,SSPN,ZIC1,ZIC4*
Cellular Function and Maintenance	4.47E-03-2.05E-02	*ACTR2,ARHGEF26,ATG3,BICD1,BOC,CDH13,CDH8,IFT27,ITPR2, PECAM1,PFN2,PLA2G5,RYR2,SPRED2*
Cellular Growth and Proliferation	4.47E-03-1.95E-02	*AGTR1,ARHGEF26,BOC,CDH13,LHX6,PECAM1,PFN2,SLC17A6,ZIC1*
Cellular Movement	4.47E-03-2.22E-02	*CDH13,DAB1,LHX6,PECAM1*
Drug Metabolism	4.47E-03-1.78E-02	*PFN2,PLA2G2A,PLA2G5,PTGS1,PVALB,SLC17A6,SSPN*
Nucleic Acid Metabolism	4.47E-03-4.71E-03	*AQP9,CPT1A,PFN2,SLC10A1*
Protein Trafficking	4.47E-03-4.47E-03	*PPP2R2B*
RNA Post-Transcriptional Modification	4.47E-03-4.47E-03	*DDX5*
Energy Production	8.92E-03-1.78E-02	*PTGS1*
Free Radical Scavenging	8.92E-03-1.34E-02	*AGTR1,PTGS1*
Gene Expression	8.92E-03-1.95E-02	*DDX5,EIF5,IGHMBP2*

RNA Damage and Repair	8.92E-03-8.92E-03	*EIF5*
Cellular Compromise	1.34E-02-1.34E-02	*AGTR1,PLA2G2A*
DNA Replication, Recombination, and Repair	1.34E-02-1.34E-02	*DHX36*
Protein Synthesis	1.34E-02-2.2E-02	*AGTR1,DDX6,DHX36,EIF5,MRPL21,MRRF,NDUFA8,NDUFS8,PECAM1, PFN2*
Vitamin and Mineral Metabolism	1.34E-02-2.22E-02	*ITPR2,PVALB*
Post-Translational Modification	1.88E-02-1.88E-02	*NDUFA8,NDUFS8*
Amino Acid Metabolism	2.22E-02-2.22E-02	*SLC17A6*

Table 4.S7 Uniquely common biological functions for carcass merit traits in a multibreed population of beef cattle

Trait	Number of common biological function	Common biological function
AFAT, CMAR, HCW, LMY, REA	22	Gene Expression, Vitamin and Mineral Metabolism, Cellular Assembly and Organization, Cellular Function and Maintenance, Cell Death and Survival, Cell-To-Cell Signaling and Interaction, Cellular Development, Cell Cycle, Cellular Movement, Cellular Growth and Proliferation, Nucleic Acid Metabolism, Post-Translational Modification, Carbohydrate Metabolism, Molecular Transport, Energy Production, Cellular Compromise, Small Molecule Biochemistry, Cell Signaling, Cell Morphology, Lipid Metabolism, Drug Metabolism, Amino Acid Metabolism
CMAR, HCW, LMY, REA	1	Protein Synthesis
AFAT, CMAR, LMY, REA	1	DNA Replication, Recombination, and Repair
AFAT, CMAR, HCW	1	Protein Trafficking
CMAR, HCW	2	RNA Post-Transcriptional Modification

Table 4.S8 The comparison of candidate genes between the current study and Wang et al

Trait	Number of candidate genes identified by Wang et al.	Number of candidate genes identified in this study	Number of overlapping genes	overlapping genes / total identified genes (%)	Overlapping gene
HCW	319	103	36	34.95%	*RNASE1, RNASE6, RNASE4, ANG2, CDH13, SULT1E1, FRAS1, ANXA3, NTRK2, AGTR1, KLHL29, SNORA70, TMEM132E, HMGCLL1, GFRAL, TRAM2, TMEM14A, GSTA2, GPR139, UMOD, PDILT, ACSM5, UQCRC2, PDZD9, VWA3A, PPRC1, SNORD22, RBM20, MAN2A1, EDEM1, KLHL31, SLC28A3, TCIRG1, CHKA, KMT5B, PERP*
REA	575	160	45	28.13%	*UMODL1, L3HYPDH, JKAMP, RTN1, LUZP2, MYO16, TRIM24, SVOPL, ATP6V0A4, GPR149, ARHGEF26, HLTF, GYG1, ZIC4, OR1J2, RYR2, RBM47, APBB2, CHST8, KCTD15, NTN1, RIMS1, KCNJ2, KCNJ16, DNER, RAB3C, CNDP2, CYB5A, PHF21B, PLSCR1, AQP9, NEDD4, TMEM171, CD247, MACF1, NPFFR2, SGCD, RNASE1, CDH13, TRAF3, AMN, CDC42BPB, FRAS1, NTRK2, COL12A1*
AFAT	189	83	23	27.71%	*CDH13, PPP2R2B, GRIK4, GRAMD1B, RGL1, CCDC92, DNAH10, MCTP2, XPC, TMEM43, CHCHD4, LUZP2, PLIN2, HAUS6, DOCK5, INPP4B, RARB, PGM5, TMEM252, THRB, FSTL4, SEC24A, TJP2*
LMY	329	43	18	41.86%	*PRTG, PYGO1, CUX2, NOS1, FOXN4, ACACB, CD247, POU2F1, NPFFR2, SGCD, UMODL1, L3HYPDH, JKAMP, PPP2R5E, STOX2, IRF2, RAB38, LUZP2*
CMAR	198	109	25	22.94%	*HLTF, GYG1, ZIC1, RYR2, RBM47, NSUN7, APBB2, RAB28, NKX3-2, BOD1L1, CDH13, SLC17A6, ITPR2, AQP9, PHLDB1, TREH, DDX6, EIF5, MARK3, CDH8, POLG2, DDX5, CEP95, LRP5, PERP*

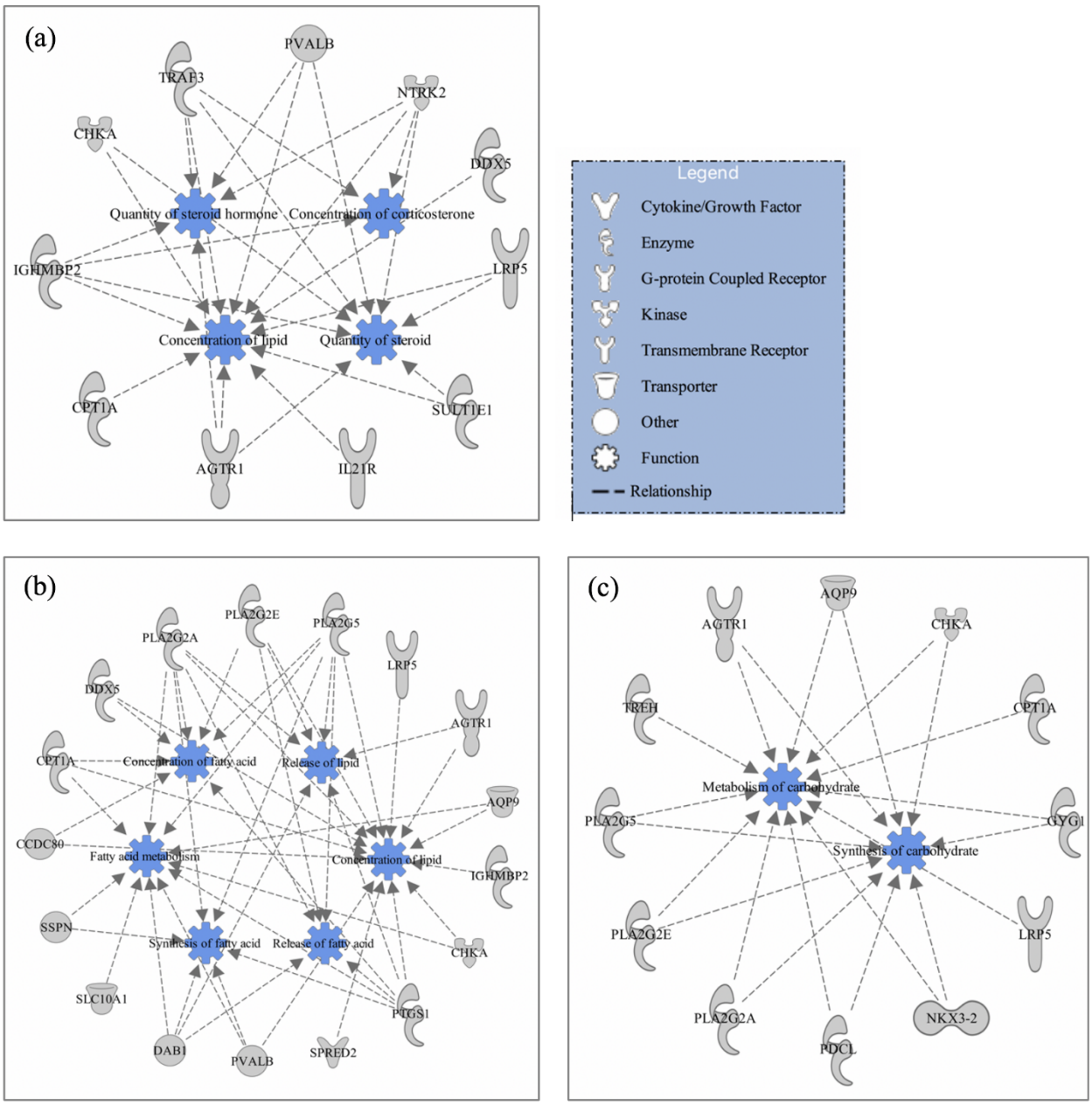

Figure 4.1 Gene networks for hot carcass weight and carcass marbling score
(a) gene network of molecular transport for hot carcass weight (HCW); **(b)** gene network of lipid metabolism for carcass marbling score (CMAR); **(c)** gene network of carbohydrate metabolism for carcass marbling score (CMAR).

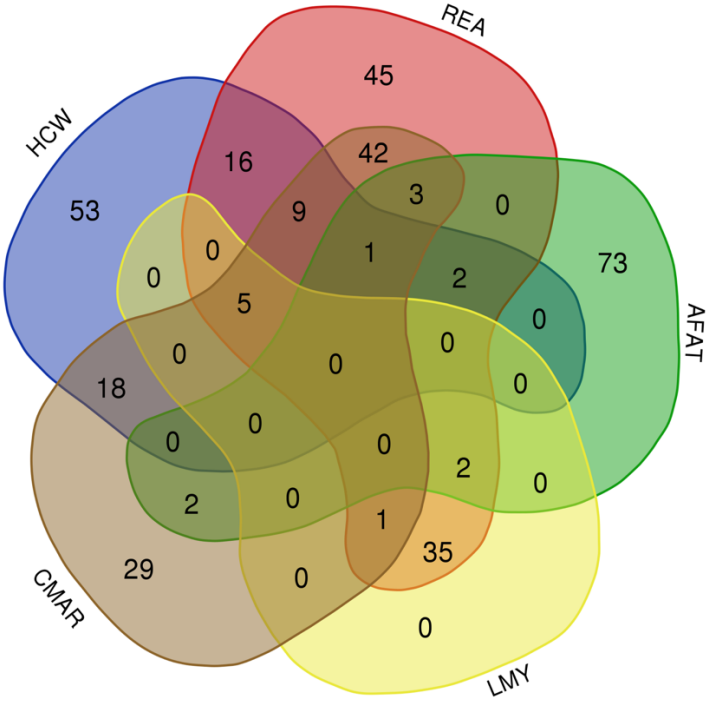

Figure 4.S1 Uniquely common candidate genes for carcass merit traits in a beef cattle multibreed population

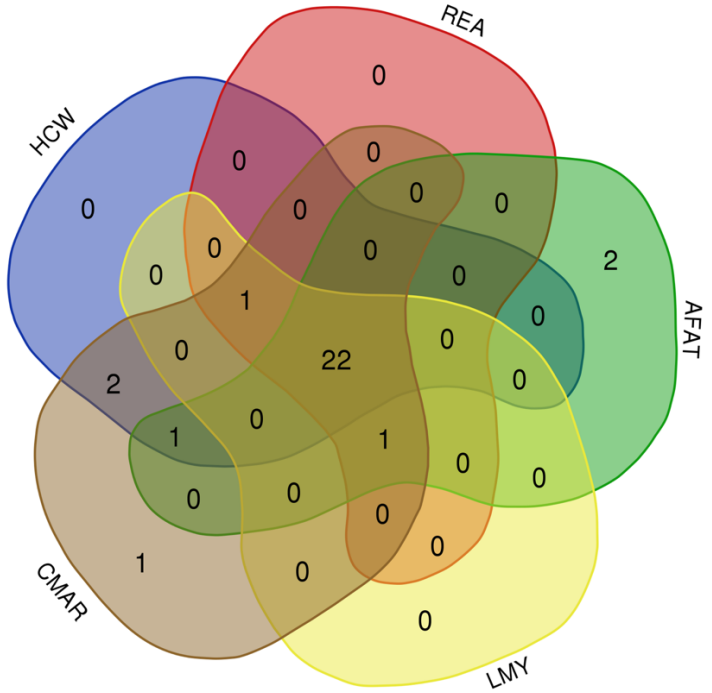

Figure 4.S2 Uniquely common biological functions for carcass merit traits in a beef cattle multibreed popular

Chapter 5. Genetic insights of bovine respiratory disease infection in feedlot crossbred cattle

5.1 Abstract

Bovine respiratory disease (BRD) is the most common and costly infectious disease affecting the well-being and productivity of beef cattle in North America. BRD is a complex disease whose development in the animals is dependent on environmental factors and host themselves (immune response and genetics). Understanding the genetic or molecular mechanisms underlying BRD infection would augment development of accurate diagnostic tools and better genetic tools that could be applied to reduce BRD prevalence, and to minimize its detrimental impact for feedlot beef production. The current study aimed to identify DNA markers associated with BRD susceptibility and universal gene expression pattern associated with BRD infection in feedlot cattle. We further investigated the association between DNA markers and gene expression to identify expression quantitative trait loci (eQTLs) associated with gene expression.

A total of 143 blood samples (80 BRD; 63 non-BRD animals) from feedlot cattle were collected for extraction of RNA and DNA. A genome-wide association study (GWAS) was performed for BRD susceptibility using 207,038 SNPs from the bovine 100K SNP array and RNA sequencing (RNA-Seq) SNP calling. Two SNPs (BovineNovelSNP1874 on chromosome 5 and BovineHD1800016801 on chromosome 18) were significantly (P-value < 1×10^{-5}) associated with BRD susceptibility. Whole blood gene expression profiles were generated for each animal using RNA-Seq to determine differential gene expression between BRD and non-BRD animals. Differentially expressed genes (DE genes) were further analyzed using functional enrichment analysis. At the significant threshold used (log2FC > 2, logCPM > 2, and FDR < 0.01), 101 DE genes were identified as associated with BRD infection and found to be involved in several

significantly (*P*-value < 0.05) enriched disease-related functions such as inflammatory response, organismal injury and abnormalities, infectious diseases, respiratory diseases and antimicrobial response. We propose to test 18 DE genes in an independent population for their potential as diagnostic biomarkers for the feedlot industry. Expression quantitative trait locus (eQTL) analysis was conducted to identify cis- and trans-eQTLs associated with DE genes. The eQTL analysis identified 420 cis-eQTLs and 144 trans-eQTLs significantly (FDR < 0.05) associated with the expression of DE genes. Additionally, the most significant SNP (BovineNovelSNP1874) identified in GWAS was a cis-eQTL for the DE gene *GPR84*. This finding suggests that this SNP could be causally associated with BRD susceptibility.

In general, the integrative analyses shed light on the understanding of genetic influences on BRD susceptibility and the molecular mechanisms underlying bovine immune responses to respiratory disease. The findings are useful for the development of a genomic selection strategy for BRD susceptibility, and for the development of new diagnostic and therapeutic tools.

5.2 Introduction

Bovine respiratory disease (BRD) is a worldwide infectious disease affecting the cattle industry. It causes a large economic burden through increased production costs associated with prevention and treatment, reduced carcass value, increased labour expense and impaired growth of animals (Griffin, 1997; Smith, 2000; Irsik et al., 2006; Schneider et al., 2010). In North America, BRD has been identified as the most expensive infectious disease of beef cattle (Taylor et al., 2010) causing high morbidity rates that can reach up to 80%, and moderate to high mortality in some feedlots (Smith, 1998; Baptista et al., 2017). Cattle in the feedlot are highly susceptible to BRD due to compromised immunity from a number of stressors including long-distance transportation and commingling cattle from different sources especially in auction markets (Taylor et al., 2010).

These stressors expose the animals to multiple BRD pathogens and provide conducive environment for emergence of opportunistic viral and bacterial infections of the respiratory tract (Griffin et al., 2010; Taylor et al., 2010; Kirchhoff et al., 2014). Early diagnosis and appropriate treatment of infected animals could enhance faster recovery from infections and potentially reduce the negative impact of the disease on animal performance and productivity. However, most clinical signs of BRD are subjective, difficult to standardize, and non-specific for BRD, which makes the diagnosis of BRD troublesome. In feedlots, the most commonly used treatment for BRD is a prophylaxis approach where animals are treated with a wide range of antibiotics before or on entry into the feedlots (Ives and Richeson, 2015), however this may lead to the development of antimicrobial resistance which is a major concern for both human and animal health (Klima et al., 2014; Stanford et al., 2020).

Investigations into the genetic architecture of resistance or susceptibility to BRD in beef cattle populations is an ongoing endeavour. Recent studies have reported low to moderate heritability in the range of 0.07 to 0.29 (Snowder et al., 2005; Schneider et al., 2010; Neibergs et al., 2014a, 2014b), which indicates potential to breed BRD resistant animals through selective breeding that will lead to a sustainable reduction in BRD incidences and potential antimicrobial resistance (Neibergs et al., 2014b; Hoff et al., 2019). Several SNPs and quantitative trait loci (QTLs) have been reported as significantly associated with BRD through genome-wide association studies (GWAS) (Neibergs et al., 2014b; Hoff et al., 2019). In addition, RNA sequencing (RNA-Seq) offers high resolution profiling of transcriptomes of individual animals in a given sample/tissue, hence allowing the discovery of transcriptome-wide expression differences between animals with contrasting phenotypes of interest (e.g., BRD and non-BRD) (Costa-Silva et al., 2017; Hrdlickova et al., 2017). Such differential gene expression between BRD and non-BRD animals could reveal

the host response to BRD infection and could help to identify potential biomarkers that could be used for BRD diagnosis. Several transcriptomic studies have been conducted to investigate the gene expression differences and host response to BRD infection (Tizioto et al., 2015; Scott et al., 2020; Sun et al., 2020; Jiminez et al., 2021). These studies have shown that host animals may regulate the immune response to defend against BRD pathogens or respond to the damage caused by BRD pathogens by influencing the expression of certain genes (i.e., differentially expressed (DE) genes), and suggest gene expression may vary with different viral and bacterial pathogen infections as well as the stage of disease development. However, these analyses only focused on the correlation between the transcriptomic level and BRD but overlooked the potential interconnections between different omics layers. Changes in gene expression are not only associated with the disease, but are also affected by genetic regulation (i.e., expression QTL, eQTL) (Cookson et al., 2009). Integration of GWAS, differential gene expression analysis and eQTL analysis could aid in interpreting the results of GWAS and identifying functional or causal SNPs. It also could provide additional insights into a probable biological basis for the disease associations and could help to identify networks of genes involved in disease pathogenesis.

Therefore, the objective of this study was to identify SNPs associated with BRD susceptibility and universal gene expression pattern associated with BRD infection in feedlots. Besides, this study investigated the association between DNA markers and gene expression to identify cis- and trans-eQTLs associated with gene expression. To achieve these objectives, we performed GWAS, differential gene expression analysis, and eQTL analysis to study the association between genotype, gene expression, and phenotype. This multi-omics study is expected to shed light on the understanding of genetic and molecular architecture of BRD in

feedlot crossbred cattle and to identify the transcriptomic biomarkers that could identify infected animals in the general beef cattle population in Canada.

5.3 Material and Methods

5.3.1 Animal population and phenotype collection

This study was conducted in accordance with the Canadian Council of Animal Care (2009) guidelines and recommendations (CCAC, 2009). All experimental procedures were reviewed and approved by the University of Calgary Veterinary Sciences Animal Care Committee (AC15-0109).

A total of 143 multi-breed and crossbred beef cattle were used in this study. Animals were conventionally raised cattle that included heifers (n = 87) and steers (n = 56). These animals were enrolled into the feedlot during the fall of 2015 at four commercial feedlots in Central/Southern Alberta. The on-arrival processing for cattle was previously described (Jiminez et al., 2021). Briefly, animals were weighed and received a subcutaneous injection of a long-acting macrolide (tulathromycin, Draxxin, 2.5 mg/kg, Zoetis, Kirkland, QC, Canada) and vaccinated against multiple bacterial and viral agents. They were also dewormed with a pour-on ivermectin solution and received a prostaglandin F2α analog injection to induce abortion as per standard feedlot procedure. While in the feedlots, animals were fed twice daily on a concentrate barley-based receiving/growing diet. This diet also contained 25 ppm of monensin (Rumensin 200, Elanco, Guelph, ON, Canada) and 35 ppm of chlortetracycline (Aureomycin 220, Zoetis). Cattle received a growth implant and second vaccination against infectious viruses at approximately 30 days after arrival. Within 50 days on feedlots, the animals were monitored daily and those that showed BRD signs and symptoms (depression, nasal or ocular discharge, cough, tachypnea, or dyspnea) were clinically examined by an experienced veterinarian. The details of clinical examinations and case definition (i.e., BRD or non-BRD cattle) were previously described (Jiminez et al., 2021). Briefly,

animals were retrospectively identified as BRD positive based on clinical examination and serum haptoglobin concentration. The animal was confirmed as BRD positive by an experienced veterinarian if the animal displayed at least one visual BRD symptom, had a rectal temperature ≥ 40 °C, abnormal lung sounds detected at auscultation, a serum haptoglobin concentration ≥ 0.25 g/L, and had no prior treatment against BRD or other diseases during the feeding period (i.e., first BRD occurrence). Blood samples were collected from the animal by jugular vein puncture for genotyping and whole-blood transcriptome using 14-gauge needles attached to a vacutainer (Tempus tubes; Thermo Fisher Scientific, ON). Once collected in Tempus tubes, samples were mixed by 20 inversions, and stored on ice until ultimately stored at -20 °C. All materials are sterile or sanitized prior to use. Meanwhile, blood samples were collected in the same way from one or two healthy matched-control pen mates (i.e., non-BRD cattle) which had no visual signs of BRD or other disease, a rectal temperature < 40 °C, no abnormal lung sounds detected at auscultation and a serum haptoglobin concentration < 0.25 g/L. The non-BRD cattle did not become positive later (i.e., remain healthy in feedlots). After sample collection, animals identified as BRD positive received an antibiotic treatment intramuscularly in combination with non-steroidal anti-inflammatory drugs, in accordance with feedlot treatment protocols.

5.3.2 RNA isolation, cDNA library preparation and sequencing

Total RNA was extracted in two batches (Batch 1, n = 47; Batch 2, n = 96). The following similar procedures were performed on all samples in both batches. Initially, total RNA was isolated from blood using a Preserved Blood RNA Purification Kit (Norgen Biotek Corp, Thorold, ON, Canada), and the quality of RNA was measured using the 2200 RNA ScreenTape TapeStation System (Agilent Technologies Inc., Cedar Creek, TX, United States) producing RNA integrity numbers (RIN) ranging from 8.0 to 9.8. Thereafter, cDNA library for each individual animal was

prepared from the extracted high-quality RNA using the TruSeq RNA Library Preparation kit v2 (Illumina, San Diego, CA, United States) and the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England Biolabs Ltd., Whitby, ON, Canada) to prepare the libraries for sequencing. Samples in both batches used the stranded library preparation process. Paired-end Sequencing was performed using the Hiseq 4000 platform and Novaseq 6000 for batch 1 and batch 2 samples respectively to generate paired-end sequences of 100 bp read length. Sequencing of samples in the two batches was performed at McGill University and Genome Quebec Innovation Center (Montreal, QC, Canada). Finally, the raw reads of 143 samples (80 BRD; 63 non-BRD) were obtained and used for downstream analyses.

5.3.3 Sequence data processing, alignment and counting

Raw reads for each sample were assessed for quality using FastQC (v0.11.8) (Andrews, 2010). The bases with low quality score (Phred quality score < 20) and 3' adapter sequences on raw reads were removed using Trimmomatic (v0.39) (Bolger et al., 2014). These cleaned-up sequences were aligned to the *Bos taurus* reference genome (ARS-UCD1.2.98, downloaded from Ensembl genome bowser) using a short read alignment software STAR (v2.7.1a) with paired-end default parameters (Dobin et al., 2013). FeatureCounts (SubRead v1.6.4) was used to count the reads that aligned to a particular annotated gene in the bovine genome (Liao et al., 2014) and these counts were consequently used for differential gene expression analysis between BRD and non-BRD animals.

5.3.4 DNA extraction and genotyping

DNA was extracted from blood samples of the transcriptome-profiled animals. The extracted DNA was used to genotype each animal using Illumina's GGP Bovine 100K microarray SNP chip (Illumina, San Diego, CA, United States). The SNPs on sex chromosomes and had minor

allele frequency < 5%, missing allele rate > 10% and failed to pass the Hardy-Weinberg equilibrium test (*P*-value < 0.0001) were excluded. The remaining SNPs (i.e.,85,100 SNPs) were used to predict breed composition using ADMIXTURE software (v1.3.0) (Alexander and Lange, 2011). The value of $K = 3$ was obtained because it had the smallest cross-validation error and yielded the most accurate breed composition prediction based on prior knowledge of breed composition on a subset of animals.

Additionally, we called additional SNP markers and genotypes from the RNA-Seq data of each animal using the Genome Analysis Toolkit best practices (GATK v3.8; Van der Auwera and O'Connor, 2020). Prior to the variant calling processes, the mapped reads from two-pass STAR alignment were sorted, had read groups added, and duplicates identified using the Picard tools package (v2.20.6). A series of processing steps including splitting "N" cigar reads (i.e., splice junction reads), reassigning mapping quality score, and base quality score recalibration were performed to improve variants calling accuracy using GATK. After data preprocessing, variants were called using the HaplotypeCaller algorithm in Genomic Variant Call Format (GVCF) mode, which included two steps: (i) variants were called individually on each sample, generating one GVCF file per sample that lists genotype likelihoods and their genome annotations; (ii) variants were called from the GVCF file through a joint genotyping analysis. The joint genotyping method is more flexible and technically easier, and is recommend for variant calling in RNA-Seq experiments (Poplin et al., 2017; Brouard et al., 2019). Stringent filtering procedures were applied to variants using the GATK Variant Filtration tool and VCFtools (v0.1.14) (Danecek et al., 2011). Indels, non-biallelic SNPs and SNPs on sex chromosomes were excluded. Then SNPs with QD < 3.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, SOR > 3.0, minor allele frequency < 5%, missing allele rate > 10% and severe departure from Hardy-Weinberg

equilibrium (*P*-value < 0.0001) were removed. Some SNPs identified in this study were not annotated in the current reference genome. These SNPs are identified with the classification "BovineNovelSNP" followed by a number in the data.

Finally, two SNP datasets (genotype derived vs. RNAseq derived) were merged based on the position of SNPs on the chromosome using Plink (v1.90b6.7) (Chang et al., 2015). For the overlapping SNPs in two SNP datasets, we retained SNPs derived from genotype. A total of 207,038 SNPs for 138 animals were obtained and used in GWAS and eQTL analysis.

5.3.5 Genome-wide association analysis for BRD susceptibility

Prior to performing the GWAS for BRD susceptibility, the phenotype of BRD was first fitted into a logistic model with a fixed effect of "feedlot", a covariate of "days on feed" and a covariate of "genomic breed composition". The fixed effect and covariates were evaluated as significant in linear regression model before fitting the logistic regression. Next, the GWAS between SNP marker genotypes (from SNP chip and RNA-seq data) and adjusted BRD status was performed using the single SNP-based mixed linear model association (mlma), as implemented in the GCTA package (v 1.93.2) (Yang et al., 2011). The linear mixed model can be described as follows:

$$y_{ij} = \mu + b_j x_{ij} + a_{ij} + e_{ij}$$

where y_{ij} is the adjusted phenotypic value of the ith animal with the jth SNP (i.e. the ijth animal), b_j is the allele substitution effect of the jth SNP, x_{ij} is the jth SNP genotype of animal i coded as 0, 1, 2 for genotypes A_1A_1, A_1A_2, and A_2A_2, respectively, a_{ij} is the additive polygenic effect of the ijth animal $\sim N(0, \boldsymbol{G}\sigma_a^2)$, and e_{ij} is the random residual effect $\sim N(0, \boldsymbol{I}\sigma_e^2)$. The genomic relationship matrix $\boldsymbol{G}$ that was derived based on total filtered SNP markers (207,038 SNPs) as described by Yang et al. (2014), which is essentially the same as the second VanRaden's

formulation (VanRaden, 2008). The SNP allele substitution effect was estimated and the significance test of the SNP allele substitution effect was conducted via a generalized least square F-test as implemented in the GCTA package. The phenotypic variance explained by each significant SNP was calculated by $\frac{2pq\beta^2}{S^2} * 100\%$, where p and q denote the SNP allele frequency of A_1 and A_2, respectively; β is the SNP allele substitution effect; $2pq\beta^2$ is the additive variance of the SNP, and S^2 is the phenotypic variance. Additionally, the variance components were estimated via a restricted maximum likelihood (REML) as implemented in the GCTA package. The genomic heritability of BRD susceptibility was calculated as a ratio of the total additive genetic variance over the phenotypic variance.

The SNPs with *P*-value < 1×10^{-5} were considered as significantly associated with BRD susceptibility according to the recommendation of The Wellcome Trust Case Control Consortium (Burton et al., 2007). The quantile-quantile (Q-Q) plot is a graphical representation of the deviation of the observed *P*-values from the null hypothesis. The Manhattan plot can visually show associations between SNPs and BRD susceptibility at the level of each chromosome. Both Q-Q plot and Manhattan plot were completed using the R package qqman (Turner, 2014).

5.3.6 Differential gene expression analysis and functional enrichment analysis

We first performed principal component analysis (PCA) based on all expressed genes of each animal to study the relationship between gene expression of animals and four feedlots. We found that animals from different feedlots were not distinct in the PCA plot containing principal components 1 and 2 (Figure 5.S1). Additionally, in this study, we aimed to determine the universal gene expression pattern of BRD. This PCA analysis illustrates that it is feasible to combine all animals from different feedlots in the differential expression analysis.

For the differential gene expression analysis, we first performed this analysis using all animals from the four feedlots. The detail of this analysis is as follows. Differential gene expression analysis between BRD and non-BRD animals was performed using the read counts and the R Bioconductor package edgeR (McCarthy et al., 2012). Lowly expressed (count per million or CPM < 0.5 in at least 63 samples) genes were filtered out from the analysis. Counts of the retained genes were then normalized using the trimmed mean M values (TMM) method (Robinson and Oshlack, 2010), to account for the technical variations between samples that may have been caused by the RNA extraction, cDNA library construction, and differences in sequencing depth (Robinson and Oshlack, 2010). The normalized counts were then modeled for differential gene expression between BRD and non-BRD animals using generalized linear models (GLM) that considered "feedlot", "sequencing batch" and "genomic breed composition". To test for significance of differential expression of a gene between the animal groups, a likelihood ratio test under negative binomial distribution assumption was performed, and those genes with Benjamini-Hochberg false discovery rate (FDR) < 0.01, log fold change (log2FC) > 2, and log counts per million (logCPM) > 2 were identified as significant differentially expressed genes between BRD and non-BRD animals. After the differential gene expression analysis for all animals, we also performed such differential expressed gene analysis for animals from each feedlot to identify the common and unique DE genes in these four feedlots. Of note, in this study, the non-BRD animals were set as the reference to indicate the DE genes that were up-regulated or down-regulated in BRD animals.

Functional enrichment for the DE genes was performed using the Ingenuity Pathway Analysis software (IPA; www.Ingenuity.com) using the Ensembl gene ID and log fold change as the inputs. In this study, biological functions were considered significantly enriched if the *P*-value

for the overlap comparison test between the input gene list and the knowledge base of IPA for a given biological function was less than 0.05.

Log transformed counts for all the DE genes for each sample were utilized for partial least squares discriminant analysis (PLS-DA) using Metaboanalyst 5.0, a web-based platform (https://www.metaboanalyst.ca/) (Pang et al., 2021). Initially, a PLS-DA model involving all the DE genes was implemented to identify the DE genes that importantly contribute to distinction of BRD from non-BRD animals. Consequently, variable importance in projection (VIP) score as a measure of informativeness of each of the DE genes in the model was calculated. In this study, we used VIP value > 1.2 as the cut-off to identify significant compounds that drives the separation of animals (i.e., potential biomarkers). A permutation test with 2,000 random resamplings was implemented to validate the reliability of the PLS-DA model.

5.3.7 eQTL analysis and eQTL annotation

We further performed eQTL analysis to identify association between expression of differentially expressed genes and SNP genotypes. Log transformed normalized counts (log_2CPM) values of DE genes on autosomes and 207,038 SNPs from 138 animals were used in eQTL analysis. The analysis of linear model was fitted to test the association of each single gene's expression and genotype classes of a SNP implemented in the R package MatrixEQTL (Shabalin, 2012). "Feedlot", "sequencing batch" and "genomic breed composition of animals" were also fitted in the model to correct for any variability in gene expression that could have been due to these factors. SNPs located within 1 Mbp around the gene transcription starting site (TSS) were tested for cis-associations, while SNPs located further than 1 Mbp or on other chromosomes were tested for trans-associations. Only those associations with FDR < 0.05 were considered significant cis- or trans-eQTLs. The significant eQTLs were then annotated as located in the TSS-promoter, exonic,

intronic, transcription termination site (TTS) or intergenic regions using the annotatePeaks.pl script from the HOMER software (http://homer.ucsd.edu/homer/ngs/annotation.html).

5.4 Results

5.4.1 Genomic background of BRD susceptibility

In this study, the genomic heritability estimate of BRD susceptibility was 0.43 ± 0.51. The GWAS results showed polygenic background of susceptibility to BRD infection in beef cattle, with only two SNPs (BovineNovelSNP1874 and BovineHD1800016801) showing significant (*P*-value > 1×10^{-5}) association with the trait (Table 5.1; Figure 5.S2 – 5.S3). The significant SNPs BovineNovelSNP1874 and BovineHD1800016801 are located on chromosome 5 and 18, respectively. Both SNPs are located in exonic regions of genes, i.e., BovineNovelSNP1874 is located in *SMUG1* while BovineHD1800016801 is located in *IGLON5*.

5.4.2 Transcriptomic architecture of BRD infection in feedlots

At the significant threshold of log fold change (log2FC) > 2, log counts per million (logCPM) > 2 and Benjamini-Hochberg false discovery rate (FDR) < 0.01, 101 genes were identified as differentially expressed between BRD and non-BRD animals, of which 7 and 94 were respectively downregulated and upregulated in the infected animals (Figure 5.1). The full list of all DE genes with related description and statistics is provided in Table 5.2. Our result showed that *interleukin 3 receptor subunit alpha* (*IL3RA)* was the most significant (FDR = 6.6×10^{-81}) upregulated gene, whereas *hemoglobin subunit beta* (*HBB*) was the most significant (FDR = 1.25×10^{-24}) downregulated gene (Table 5.2). In terms of fold change, *leucine rich alpha-2-glycoprotein 1* (*LRG1*) showed the highest fold change (FC = 162.87), and *hemoglobin subunit alpha 1* (*HBA1*) showed the lowest fold change (FC = 0.11) (Table 5.2).

Of the 101 DE genes, 88 successfully mapped to the IPA database for functional enrichment analysis. The DE genes were significantly (*P*-value < 0.05) involved in 17 disease-related biological functions of which inflammatory response was the most significant with 60 DE genes. The top 10 most enriched functions are presented in Table 5.3, while all the 17 functions are presented in the Table 5.S1. Within the inflammatory response function, 3 DE genes (*ARG1*, *ALOX15*, and *ALAS2*) were downregulated, and 57 DE genes (e.g., *IL3RA*, *LRG1*, *BPI*, *CFB*, *GPR84*, *MMP9*, and *CA4*) were upregulated in the BRD animals. Furthermore, within the inflammatory response function, enriched innate immune response related processes such as leukocyte immune response, activation and migration of macrophages and neutrophils, and antimicrobial response were predicted to be activated or upregulated in the BRD animals (Figure 5.2). Adaptive immune response related processes such as activation of antigen processing cells, and cellular immune response were identified as enriched and predicted to be activated in the BRD animals. Some of the key DE genes as demonstrated by their involvement in numerous immune functions included *LCN2*, *S100A8*, *S100A9*, *S100A12*, *LTF*, *IL12B*, *CHI3L1* and *DEFB4A* (Figure 5.2).

Additionally, at the same threshold, we also identified DE genes using animals separately for each feedlot: 127, 69, 50 and 21 DE genes were identified for the different feedlots (Table 5.S2 – S5). The common and unique DE genes are summarized in the Table 5.S6. The results showed that 12 DE genes were identified in all feedlots, 32 DE genes were identified in three feedlots, and 31 DE genes were identified in two feedlots. Interestingly, *LRG1* was the DE gene with highest fold change in all feedlots. *HBB*, *HBA* and *HBA1* were downregulated DE genes in three feedlots. We also compared the DE genes identified using the combined animals (i.e., all animals) and those animals from each feedlot. The result showed 96 genes were overlapped (Table 5.S7).

Partial least squares discriminant analysis (PLS-DA) revealed clear distinction between BRD infected and the non-infected animals as shown in Figure 5.3. Of the 101 DE genes used in the PLS-DA model, 18 DE genes had Variable Importance in the Projection (VIP) scores greater than 1.2 (Figure 5.4) and *LRG1* had the highest VIP values. The result indicated the suitability of these 18 DE genes as biomarkers for identifying BRD infected and non-infected animals.

5.4.3 Gene expression and genotype associations

At FDR < 0.05, we identified 420 cis-eQTLs and 144 trans-eQTLs associated with the expression of DE genes (Table 5.S8 – 5.S9). Some cis-eQTLs and trans-eQTLs were associated with more than one DE gene associated with BRD. For example, the SNP BovineNovelSNP2890 was a cis-eQTL associated with the expression of the DE gene *BST1* and a trans-eQTL associated with the expression of other 6 DE genes (*GPR84*, *NUPR1*, *ART5*, *CFB*, *SLC6A2*, and *ADGRE1*). Similarly, the expression of a DE gene could also be associated with more than one cis- or trans-eQTLs. Of note, the eQTL analysis showed that the SNP (BovineNovelSNP1874) with the smallest *P*-value in GWAS (Table 5.1) was a cis-eQTL associated with the expression of the DE genes *GPR84* (Table 5.S8 – 5.S9). Additionally, the eQTL annotation showed that the eQTL SNPs identified in this study are mostly located in the intronic and exonic regions (Figure 5.5).

5.5 Discussion

The BRD and non-BRD animals used in the current study were fed in feedlots with the same operations and similar environments. It is expected that all animals in the study were equally exposed to BRD causing pathogens, therefore, all BRD animals are assumed to be susceptible while non-BRD animals are resistant. Disease susceptibility and resistance was defined in relation to BRD in general and not according to specific pathogens involved. Based on this assumption, heritability of BRD susceptibility was estimated and SNPs associated with BRD susceptibility in

beef cattle were identified through GWAS. A moderate heritability estimate of 0.43 ± 0.51 for BRD susceptibility was observed in the studied population. This estimate had a relatively large standard error due to limited sample size (n = 138). Previous studies have reported low to moderate heritability for BRD susceptibility that ranged from 0.07 ± 0.01 to 0.29 ± NA in different cattle populations (Snowder et al., 2005; Schneider et al., 2010; Neibergs et al., 2014a, 2014b). For GWAS analysis, two SNPs: BovineNovelSNP1874 and BovineHD1800016801 that were associated with BRD susceptibility were identified (Table 5.1). Interestingly, the most significant SNP (BovineNovelSNP1874) explained 17% of the phenotypic variance for BRD susceptibility. This implies that this SNP could be a major quantitative trait nucleotide (QTN) or in linkage disequilibrium with a major QTL for BRD susceptibility in the studied population. However, the proportion of phenotypic variance explained by significant SNPs could be overestimated because of the limited number of animals used. In addition, the low coverage depth of SNP calling and low minor allele frequency of BovineNovelSNP1874 may also lead to a false-positive result. Thus, future research utilising larger sample size is warranted to verify these results.

In addition to the genetic background, the transcriptome variability associated with BRD was investigated. A total of 101 DE genes were identified between BRD and non-BRD animals. About 93% of the DE genes were upregulated in the BRD animals (Table 5.2). Among these upregulated genes, *IL3RA* and *LRG1* showed the strongest association to BRD in terms of statistical significance and fold change, respectively. *IL3RA* encodes the protein of interleukin 3 receptor subunit alpha which is a cytokine receptor protein for the interleukin 3 (IL3), colony stimulating factor 2 (CSF2/GM-CSF) and interleukin 5 (IL5) (Milatovich et al., 1993). The cytokine IL3 is generated from T cells and stem cells, and is involved in macrophage activation and regulation of cytokine production (Frendl, 1992). On the other hand, IL-5 is produced by

CD4+ T cells and causes B-cell growth factor and differentiation, IgA selection, eosinophil activation and increased production of innate immune cells (Akdis et al., 2011). This study also identified DE genes (*IL1R2*, *IL1RAP*, and *IL12B*) related to interleukin-1 (IL-1) and interleukin-12 (IL-12), which causes lymphocyte activation, macrophage stimulation, increased leukocyte adhesion and release of acute phase proteins by the liver, or induced interferon gamma production by T cells and NK cells (Arena et al., 1998; Akdis et al., 2011; Dinarello, 2018; Jiang et al., 2018). The *LRG1* is an important DE gene that has also been reported as associated with BRD in previous transcriptomic studies (Tizioto et al., 2015; Scott et al., 2020; Jiminez et al., 2021). *LRG1* encodes the protein of leucine rich alpha-2-glycoprotein 1 that has been reported to be packaged into the granule compartment of human neutrophils and secreted upon neutrophil activation (Druhan et al., 2017). For downregulated genes, the top 3 genes (*HBA1*, *HBA*, and *HBB*) are all related to hemoglobin - the oxygen-carrying protein within red blood cells. Specifically, *HBA1* and *HBA* encodes α-globins, and *HBB* encodes β-globins, which are the two main globins that compose hemoglobin (Marengo-Rowe, 2006). Thus, the low expressed level of *HBA1*, *HBA* and *HBB* may indicate a low hemoglobin count in infected cattle (BRD susceptible cattle) associated with anemia. Additionally, the low hemoglobin count may be associated with elevated iron levels. Iron is important to pathogen growth, and elevated iron may lead to a higher chance of pathogen infection. Future studies should determine iron levels to investigate this hypothesis.

Since the type and number of pathogens may be different in different feedlots, we also analyzed DE genes separately in each feedlot (Table 5.S2 – 5.S5). The results showed some genes were differentially expressed in all feedlots (e.g., 12 overlapping DE genes), which indicates that these genes may play key role in the BRD infection. Especially, the gene *LRG1* was the DE gene with the highest fold change in all feedlots (Table 5.S2 – 5.S5). We further compared the DE genes

identified in combined animals (i.e., all animals) and separated animals (animals from each feedlot), about 95% of DE genes (96 DE genes) that were identified using combined animals overlapped with those identified using separate animals (Table 5.S6). Also, since the combined population contains all animals from each feedlot, the number of false positive genes could be reduced. Therefore, the DE genes identified in the combined population could represent the transcriptome variability associated with BRD.

Some of the DE genes identified in the current study have been identified as associated with BRD in beef cattle in other similar studies but from DE analysis of lymph nodes (Tizioto et al., 2015), bronchial epithelial cells (N'jai et al., 2013), and blood (Scott et al., 2020; Jiminez et al., 2021). Compared with the results of Tizioto et al. (2015), 26, 35, 29, 39, 20 and 8 of DE genes identified in this study were common with those identified in the lymph node of animals who were challenged by BRSV, IBR, BVDV, *M. haemolytica*, *P. multocida*, and *M. bovis*, respectively (overlapping genes are shown in the Table 5.S5). In addition to identifying DE genes specific to individual challenges, Tizioto et al. (2015) found 25 genes expressed differentially in all the infections, of which 5 genes (*S100A8*, *S100A9*, *MMP9*, *TGM3*, and *PGLYRP1*) were also identified as DE genes in the current study (Table 5.S10). These genes may be differentially expressed in all pathogen challenges because they are related to innate immune cells. For example, *S100A8* and *S100A9* are expressed in neutrophils and monocytes (Edgeworth et al., 1991) and are known danger-associated molecular patterns that bind pattern recognition receptors in response to inflammation (Schiopu and Cotoi, 2013). Additionally. N'jai et al. (2013) reported the top 70 DE genes identified in bovine bronchial epithelial cells, 3 genes (*CA4*, *TNFAIP6,* and *HP*) were also reported in our study and previous studies (Tizioto et al., 2015; Jiminez et al., 2021). Comparing our results with DE genes identified in blood samples from other studies (Scott et al., 2020; Jiminez

et al., 2021), more common DE genes, such as *LRG1*, *CFB*, and *ALOX15,* were observed. Therefore, the DE genes associated with BRD in different populations have good consistency. Furthermore, BRD is a polymicrobial disease that is usually the result of co-infection of several common viral and bacterial pathogens (Dabo et al., 2008; Rice et al., 2008; Griffin et al., 2010; Klima et al., 2014). The infection of different pathogens may cause different immune responses and related gene expression of the host (N'jai et al., 2013; Tizioto et al., 2015). Through the comparison between our results and results of Tizioto et al. (2015) (Table 5.S10), the infection process in our population seems to involve multiple pathogens as well. However, the study design and the objectives of our study were to determine the common immune response to BRD infection and to identify transcriptome biomarkers that could be used in different populations and feedlots. The fact that the expression of some genes is associated with more than one pathogen and some genes could respond to all pathogen infections (N'jai et al., 2013; Tizioto et al., 2015) makes it difficult to distinguish specific pathogen infections based on gene expression alone. Therefore, future study to evaluate the influence of pathogens on gene expression is recommended. This may also help to identify pathogen-specific DE genes.

Investigation into the biological involvement of the DE genes revealed inflammatory response as the most significant enriched function. In animals, inflammatory response is a biological response of the immune system to injurious stimuli, such as pathogens, damaged cells and toxic compounds (Ferrero-Miliani et al., 2007; Medzhitov, 2010). This response is aimed at clearing the immune insulting agents and initiating healing (Ferrero-Miliani et al., 2007; Medzhitov, 2010). Upon recognition of the pathogenic agents, the immune system responds to such attack by recruiting and activating the phagocytic cells such as macrophages and neutrophils, and those phagocytes that are tasked with the immediate destruction and clearing of the pathogenic

agents from the body (Ackermann et al., 2010; Mantovani et al., 2011). Interestingly, activation and recruitment of both neutrophils and macrophages were among the processes identified as enriched within inflammatory response in the current study (Figure 5.2). These processes were predicted to be activated in the BRD animals compared to the non-BRD animals, indicating that the inflammatory response plays a key role in defense against pathogenic infection of BRD. Previous transcriptome studies of blood and other immune organs also demonstrated the significant association of inflammatory response with BRD status (N'jai et al., 2013; Scott et al., 2020). Some of the interesting genes involved in inflammatory response due to their involvement in multiple innate immune response within this function included *LCN2*, *S100A8*, *S100A9*, *S100A12*, *LTF*, *IL12B*, *CHI3L1*, *DEFB4A,* and *MMP9.* In line with our results and speculation, Tizioto et al. (2015) reported DE genes and pathways that were found to be common to all pathogen challenges were up-regulated (e.g., *S100A8*, *S100A9*, and *MMP9*) in the challenged animals (BRD animals) and appear to primarily be related to the innate immune response. Therefore, this information revealed the biological processes that DE genes were involved in, which provides insights into the biological background of BRD infection and host immune response.

Furthermore, eQTL analysis identified many associations between DNA markers and gene expression, indicating complex genetic regulation of gene expression. For example, one trans-eQTL could affect the expression of many DE genes, suggesting that there may be a trans-eQTL hotspot (Figure 5.6). This information obtained from eQTL analysis could help to understand our GWAS result and illustrate the causality between the significant SNP and BRD susceptibility. For example, the SNP (BovineNovelSNP1874) with the lowest *P*-value in GWAS showed a cis effect on the DE gene *GPR84* (Table 5.S8 – 5.S9). The SNP could indeed be a causal mutation for

variability in BRD susceptibility or resistance that needs further investigation and validation in independent populations or cell culture or transgenic assays. In addition, the results obtained from eQTL analysis could help to pinpoint causal SNPs associated with susceptibility to BRD. For example, Neibergs et al. (2014b) reported a genomic region covering *BPI* as associated with BRD susceptibility in Holstein calves, thus indicating that variants within or near this gene, have functional relevance in modulating susceptibility to BRD in cattle. *BPI* was also a DE gene associated with BRD in the current study and previous studies (Tizioto et al., 2015; Jiminez et al., 2021). *BPI* encodes the bactericidal permeability increasing protein, a critical protein involved in neutralizing gram negative bacteria lipopolysaccharide antigen and mediates and promotes gram negative bacteria recognition by monocytes for phagocytosis (Yu and Song, 2020). Through the eQTL analysis, we further identified the most likely causal SNP among all variants within or near the gene *BPI*. The SNP (rs209419196) was the most significant SNP (*P*-value < 2.1×10^{-6}, FDR < 0.006) among 6 cis-eQTLs associated with the expression of *BPI* (Figure 5.7), and the expression of *BPI* was significantly (*P*-value < 0.05) decreased as the number of "T" alleles increased in the genotype (Figure 5.8). According to eQTL annotation analysis, rs209419196 was predicted to be in the promoter region of *BPI*, which is located 92 bp downstream of the 5' end of the transcription start site for the transcript (ENSBTAT00000077785) of *BPI*. Therefore, the results of eQTLs and their annotation cannot only provide important reference information for GWAS interpretation and causal SNP identification, but also provide additional insights into potential molecular mechanism of gene expression. Also, identification of these SNP markers provides more functional information that could be utilized to enhance genomic selection for BRD resistance in beef cattle.

In view of the multiple etiology of BRD and a complex interaction among risk factors, BRD is difficult to control and prevent. Conventionally, BRD diagnosis is based on clinical signs,

and varies among premises, calf caretakers, producers, and herd veterinarians leading to a proportion of false-negative and false-positive diagnoses (Moisá et al., 2019). Such diagnostic inaccuracies lead to progression of disease, misuse of antimicrobials, production losses, and suboptimal animal welfare outcomes (Moisá et al., 2019). Therefore, accurate diagnostic methods for BRD are still needed. Blood transcriptomic biomarkers have been proposed to be used in the identification of BRD cattle in feedlots (Sun et al., 2020). In the current study, we propose 18 DE genes as potential diagnostics biomarkers for BRD in our population. The most informative marker *LRG1* has been previously identified as a potential biomarker of different infections in humans (Wu et al., 2015; Fujimoto et al., 2020; Yang et al., 2020; Ma et al., 2021). Although these biomarkers are very promising, validation in other independent beef cattle populations is required for evaluating their performance.

5.6 Conclusion

Genomic and transcriptomic tools were applied to elucidate the genetic and molecular background of BRD infection in feedlot beef cattle. Two SNPs associated with BRD susceptibility were identified through GWAS. Transcriptomic and functional analyses revealed 101 DE genes associated with BRD infection. These genes were mainly involved in inflammatory response processes such as recruitment and activation of phagocytes. The most significant SNP (BovineNovelSNP1874) from the GWAS analysis was also a cis-eQTL associated with a DE gene *GPR84*. This indicates that such integrative analysis could help to interpret previous GWAS results and identify causal SNPs associated with BRD susceptibility. Additionally, we recommend testing 18 DE genes in an independent population to investigate their potential as diagnostic biomarkers in feedlots. Therefore, these biomarkers and causal SNPs identified in our population warrants further investigation to validate their usefulness in other beef cattle populations.

5.7 References

Ackermann, M. R., Derscheid, R., and Roth, J. A. (2010). Innate Immunology of Bovine Respiratory Disease. *Vet. Clin. North Am. Food Anim. Pract.* 26, 215. doi:10.1016/J.CVFA.2010.03.001.

Akdis, M., Burgler, S., Crameri, R., Eiwegger, T., Fujita, H., Gomez, E., et al. (2011). Interleukins, from 1 to 37, and interferon-γ: Receptors, functions, and roles in diseases. *J. Allergy Clin. Immunol.* 127, 701-721.e70. doi:10.1016/j.jaci.2010.11.050.

Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12, 246. doi:10.1186/1471-2105-12-246.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Arena, W. P., Malyak, M., Guthridge, C. J., and Gabay, C. (1998). Interleukin-1 receptor antagonist: Role in biology. *Annu. Rev. Immunol.* 16, 27–55. doi:10.1146/annurev.immunol.16.1.27.

Baptista, A. L., Rezende, A. L., Fonseca, P. de A., Massi, R. P., Nogueira, G. M., Magalhães, L. Q., et al. (2017). Bovine respiratory disease complex associated mortality and morbidity rates in feedlot cattle from southeastern Brazil. *J. Infect. Dev. Ctries.* 11, 791–799. doi:10.3855/jidc.9296.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.

Brouard, J. S., Schenkel, F., Marete, A., and Bissonnette, N. (2019). The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J. Anim. Sci. Biotechnol.* 10, 44. doi:10.1186/s40104-019-0359-0.

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678. doi:10.1038/nature05911.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. doi:10.1186/S13742-015-0047-8.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* 10, 184–194. doi:10.1038/nrg2537.

Costa-Silva, J., Domingues, D., and Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* 12, e0190152. doi:10.1371/JOURNAL.PONE.0190152.

Dabo, S. M., Taylor, J. D., and Confer, A. W. (2008). Pasteurella multocida and bovine respiratory disease. *Anim. Heal. Res. Rev.* 8, 129–150. doi:10.1017/S1466252307001399.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.

doi:10.1093/BIOINFORMATICS/BTR330.

Dinarello, C. A. (2018). Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol. Rev.* 281, 8. doi:10.1111/IMR.12621.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.

Druhan, L. J., Lance, A., Li, S., Price, A. E., Emerson, J. T., Baxter, S. A., et al. (2017). Leucine Rich α-2 Glycoprotein: A Novel Neutrophil Granule Protein and Modulator of Myelopoiesis. *PLoS One* 12, e0170261. doi:10.1371/JOURNAL.PONE.0170261.

Edgeworth, J., Gorman, M., Bennett, R., Freemont, P., and Hogg, N. (1991). Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J. Biol. Chem.* 266, 7706–7713. doi:10.1016/S0021-9258(20)89506-4.

Ferrero-Miliani, L., Nielsen, O. H., Andersen, P. S., and Girardin, S. E. (2007). Chronic inflammation: Importance of NOD2 and NALP3 in interleukin-1β generation. *Clin. Exp. Immunol.* 147, 227–235. doi:10.1111/j.1365-2249.2006.03261.x.

Frendl, G. (1992). Interleukin 3: From colony-stimulating factor to pluripotent immunoregulatory cytokine. *Int. J. Immunopharmacol.* 14, 421–430. doi:10.1016/0192-0561(92)90172-H.

Fujimoto, M., Matsumoto, T., Serada, S., Tsujimura, Y., Hashimoto, S., Yasutomi, Y., et al. (2020). Leucine-rich alpha 2 glycoprotein is a new marker for active disease of tuberculosis. *Sci. Rep.* 10, 3384. doi:10.1038/s41598-020-60450-3.

Griffin, D. (1997). Economic impact associated with respiratory disease in beef cattle. *Vet. Clin. North Am. Food Anim. Pract.* 13, 367–377. doi:10.1016/S0749-0720(15)30302-9.

Griffin, D., Chengappa, M. M., Kuszak, J., and McVey, D. S. (2010). Bacterial pathogens of the bovine respiratory disease complex. *Vet. Clin. North Am. - Food Anim. Pract.* 26, 381–394. doi:10.1016/j.cvfa.2010.04.004.

Hoff, J. L., Decker, J. E., Schnabel, R. D., Seabury, C. M., Neibergs, H. L., and Taylor, J. F. (2019). QTL-mapping and genomic prediction for bovine respiratory disease in U.S. Holsteins using sequence imputation and feature selection. *BMC Genomics* 20, 555. doi:10.1186/s12864-019-5941-5.

Hrdlickova, R., Toloue, M., and Tian, B. (2017). RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip. Rev. RNA* 8, e1364. doi:10.1002/WRNA.1364.

Irsik, M., Langemeier, M., Schroeder, T., Spire, M., and Deen Roder, J. (2006). Estimating the Effects of Animal Health on the Performance of Feedlot Cattle. *Bov. Pract.* 40, 65–74. doi:org/10.21423/bovine-vol40no2p65-74.

Ives, S. E., and Richeson, J. T. (2015). Use of Antimicrobial Metaphylaxis for the Control of Bovine Respiratory Disease in High-Risk Cattle. *Vet. Clin. North Am. - Food Anim. Pract.* 31, 341–350. doi:10.1016/j.cvfa.2015.05.008.

Jiang, T., Huang, Z., Zhang, S., Zou, W., Xiang, L., Wu, X., et al. (2018). miR-23b inhibits proliferation of SMMC-7721 cells by directly targeting IL-11. *Mol. Med. Rep.* 18, 1591–

1599. doi:10.3892/MMR.2018.9151.

Jiminez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front. Genet.* 12, 257. doi:10.3389/fgene.2021.627623.

Kirchhoff, J., Uhlenbruck, S., Goris, K., Keil, G. M., and Herrler, G. (2014). Three viruses of the bovine respiratory disease complex apply different strategies to initiate infection. *Vet. Res.* 45, 20. doi:10.1186/1297-9716-45-20.

Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., et al. (2014). Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *J. Clin. Microbiol.* 52, 438–448. doi:10.1128/JCM.02485-13.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/BIOINFORMATICS/BTT656.

Ma, H., Lu, F., Guo, Y., Shen, Z., and Chen, Y. (2021). The Prognostic Value of Leucine-Rich α 2 Glycoprotein 1 in Pediatric Spinal Cord Injury. *Biomed Res. Int.* 2021, 7365204. doi:10.1155/2021/7365204.

Mantovani, A., Cassatella, M. A., Costantini, C., and Jaillon, S. (2011). Neutrophils in the activation and regulation of innate and adaptive immunity. *Nat. Rev. Immunol.* 11, 519–531. doi:10.1038/nri3024.

Marengo-Rowe, A. J. (2006). Structure-function relations of human hemoglobins. *Proc. (Bayl. Univ. Med. Cent).* 19, 239–245. doi:10.1080/08998280.2006.11928171.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288. doi:10.1093/NAR/GKS042.

Medzhitov, R. (2010). Inflammation 2010: New Adventures of an Old Flame. *Cell* 140, 771–776. doi:10.1016/j.cell.2010.03.006.

Milatovich, A., Kitamura, T., Miyajima, A., and Francke, U. (1993). Gene for the alpha-subunit of the human interleukin-3 receptor (IL3RA) localized to the X-Y pseudoautosomal region. *Am. J. Hum. Genet.* 53, 1146.

Moisá, S. J., Aly, S. S., Lehenbauer, T. W., Love, W. J., Rossitto, P. V., Eenennaam, A. L. Van, et al. (2019). Association of plasma haptoglobin concentration and other biomarkers with bovine respiratory disease status in pre-weaned dairy calves. *J. Vet. Diagn. Invest.* 31, 40. doi:10.1177/1040638718807242.

N'jai, A. U., Rivera, J., Atapattu, D. N., Owusu-Ofori, K., and Czuprynski, C. J. (2013). Gene expression profiling of bovine bronchial epithelial cells exposed in vitro to bovine herpesvirus 1 and Mannheimia haemolytica. *Vet. Immunol. Immunopathol.* 155, 182–189. doi:10.1016/j.vetimm.2013.06.012.

Neibergs, H. L., Neibergs, J. S., Wojtowicz, A. J., Taylor, J. F., Seabury, C. M., and Womack, J. E. (2014a). Economic benefits of using genetic sele ction to reduce the prev alence of bovine respiratory disease complex in beef feedlot cattle. in *Proceedings of the Beef*

Improvement Federation (BIF) Annual Meeting and Research Symposium, 82–86.

Neibergs, H. L., Seabury, C. M., Wojtowicz, A. J., Wang, Z., Scraggs, E., Kiser, J. N., et al. (2014b). Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. *BMC Genomics* 15, 1164. doi:10.1186/1471-2164-15-1164.

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., et al. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* 49, W388–W396. doi:10.1093/NAR/GKAB382.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Auwera, G. A. Van der, et al. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178. doi:10.1101/201178.

Rice, J. A., Carrasco-Medina, L., Hodgins, D. C., and Shewen, P. E. (2008). Mannheimia haemolytica and bovine respiratory disease. *Anim. Heal. Res. Rev.* 8, 117–128. doi:10.1017/S1466252307001375.

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. doi:10.1186/GB-2010-11-3-R25.

Schiopu, A., and Cotoi, O. S. (2013). S100A8 and S100A9: DAMPs at the crossroads between innate immunity, traditional risk factors, and cardiovascular disease. *Mediators Inflamm.* 2013, 828354. doi:10.1155/2013/828354.

Schneider, M. J., Tait, R. G., Ruble, M. V., Busby, W. D., and Reecy, J. M. (2010). Evaluation of fixed sources of variation and estimation of genetic parameters for incidence of bovine respiratory disease in preweaned calves and feedlot cattle. *J. Anim. Sci.* 88, 1220–1228. doi:10.2527/jas.2008-1755.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole blood transcriptomic analysis of beef cattle at arrival identifies potential predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. *PLoS One* 15, e0227507. doi:10.1371/JOURNAL.PONE.0227507.

Shabalin, A. A. (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 28, 1353–1358. doi:10.1093/BIOINFORMATICS/BTS163.

Smith, R. A. (1998). Impact of disease on feedlot performance: a review. *J. Anim. Sci.* 76, 272–274. doi:10.2527/1998.761272X.

Smith, R. A. (2000). Effects of feedlot disease on economics, production and carcass value. in *The American Association of Bovine Practitioners Proceedings*, 125–128.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2005). Influence of breed, heterozygosity, and disease incidence on estimates of variance components of respiratory disease in preweaned beef calves. *J. Anim. Sci.* 83, 1247–1261. doi:10.2527/2005.8361247x.

Stanford, K., Zaheer, R., Klima, C., McAllister, T., Peters, D., Niu, Y. D., et al. (2020). Antimicrobial resistance in members of the bacterial bovine respiratory disease complex isolated from lung tissue of cattle mortalities managed with or without the use of antimicrobials. *Microorganisms* 8, 288. doi:10.3390/microorganisms8020288.

Sun, H. Z., Srithayakumar, V., Jiminez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics* 112, 3968–3977. doi:10.1016/J.YGENO.2020.07.014.

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., and Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *Can. Vet. J.* 51, 1095–1102.

Tizioto, P. C., Kim, J. W., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015). Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: An RNA-sequence analysis of the bronchial lymph node transcriptome. *PLoS One* 10, e0131459. doi:10.1371/journal.pone.0131459.

Turner, S. D. (2014). qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*, 005165. doi:10.1101/005165.

Van der Auwera, G. A., and O'Connor, B. D. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. 1st Editio. O'Reilly Media, Inc.

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

Wu, J., Yin, H., Zhu, J., Buckanovich, R. J., Thorpe, J. D., Dai, J., et al. (2015). Validation of LRG1 as a Potential Biomarker for Detection of Epithelial Ovarian Cancer by a Blinded Study. *PLoS One* 10, e0121112. doi:10.1371/JOURNAL.PONE.0121112.

Yang, F. J., Hsieh, C. Y., Shu, K. H., Chen, I. Y., Pan, S. Y., Chuang, Y. F., et al. (2020). Plasma Leucine-Rich α-2-Glycoprotein 1 Predicts Cardiovascular Disease Risk in End-Stage Renal Disease. *Sci. Rep.* 10, 5988. doi:10.1038/s41598-020-62989-7.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi:10.1016/j.ajhg.2010.11.011.

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46, 100–106. doi:10.1038/ng.2876.

Yu, Y., and Song, G. (2020). "Lipopolysaccharide-binding protein and bactericidal/permeability-increasing protein in lipid metabolism and cardiovascular diseases," in *Lipid Transfer in Lipoprotein Metabolism and Cardiovascular Disease*, ed. X. Jiang (Springer, Singapore), 27–35. doi:10.1007/978-981-15-6082-8_3.

Table 5.1 SNPs significantly associated with BRD susceptibility

SNP	Chromosome	Position (bp)	Minor allele frequency	b	se	*P*-value
BovineNovelSNP1874	5	25858264	0.07	1.18124	0.260658	5.8496E-06
BovineHD1800016801	18	57400705	0.20	0.674002	0.150619	7.65E-06

Table 5.2 Differentially expressed genes between BRD and non-BRD animals

Gene ID	Gene name	Chr	Start (bp)	End (bp)	logFC	logCPM	*P*-Value	FDR
ENSBTAG00000054844	*HBA1*	25	219512	220318	-3.1884825	11.0003434	3.80E-18	4.95E-17
ENSBTAG00000051412	*HBA*	25	216448	217264	-3.1783435	10.998965	5.02E-18	6.47E-17
ENSBTAG00000037644	*HBB*	15	48362236	48363996	-3.0503428	11.0891232	4.35E-26	1.25E-24
ENSBTAG00000013178	*ALAS2*	X	92485200	92512358	-2.5422182	5.31393543	2.27E-20	3.70E-19
ENSBTAG00000004824	*REEP1*	11	48518871	48654804	-2.4186155	2.11319391	1.02E-14	9.21E-14
ENSBTAG00000011990	*ALOX15*	19	26697231	26705809	-2.3472557	4.91470179	1.35E-21	2.46E-20
ENSBTAG00000012403	*ARG1*	9	69522090	69539079	-2.1567384	2.58862838	7.74E-13	5.65E-12
ENSBTAG00000011465	*MYBPH*	16	957169	965458	2.00248844	6.08726325	1.16E-33	6.80E-32
ENSBTAG00000010464	*MN1*	17	67173546	67221633	2.0092144	2.52554496	4.76E-61	3.00E-58
ENSBTAG00000006156	*BST1*	6	110872406	110901795	2.01176374	5.49028329	3.76E-43	5.60E-41
ENSBTAG00000020430	*GLT1D1*	17	48170769	48286681	2.01978988	3.08211947	5.24E-43	7.63E-41
ENSBTAG00000050072		11	98685584	98686780	2.04689154	2.42678465	1.30E-24	3.23E-23
ENSBTAG00000052465		9	42992439	42996483	2.04843735	6.79225753	2.32E-37	1.97E-35
ENSBTAG00000013368	*ANKRD22*	26	10506966	10534508	2.04918903	3.86149275	5.75E-36	4.16E-34
ENSBTAG00000002148	*RAB3D*	7	15736270	15747432	2.05267427	7.19558878	4.63E-53	1.36E-50
ENSBTAG00000023648	*ART5*	15	51386896	51443955	2.05464096	7.1142583	1.13E-45	2.00E-43
ENSBTAG00000011037	*RBPMS2*	10	45313259	45339402	2.06745272	2.97256633	1.10E-39	1.16E-37
ENSBTAG00000039556	*WIPI1*	19	61752501	61782684	2.07573348	4.45488993	4.77E-48	1.05E-45
ENSBTAG00000006921	*ABCA6*	19	61421551	61482846	2.08080075	6.47238725	2.33E-49	5.61E-47

ENSBTAG00000013555	*ACVR1B*	5	27884739	27915481	2.08763819	5.89682031	6.85E-60	3.49E-57
ENSBTAG00000000783	*TGFA*	11	13861144	13975868	2.09292441	2.66737919	1.94E-36	1.47E-34
ENSBTAG00000004150	*NRG1*	27	28529266	28651519	2.09330082	4.79180584	1.59E-50	4.13E-48
ENSBTAG00000022779	*OLFM4*	12	10641503	10665121	2.09662554	3.27667994	2.97E-28	1.04E-26
ENSBTAG00000006990	*MYRF*	29	40197141	40232303	2.11379718	2.5779438	2.04E-25	5.47E-24
ENSBTAG00000037826		15	51474978	51495512	2.11535036	2.44646954	5.59E-33	3.01E-31
ENSBTAG00000004716	*RETN*	7	16513147	16514562	2.13295268	5.03641985	1.76E-19	2.63E-18
ENSBTAG00000014046	*BPI*	13	67172370	67201998	2.13491623	5.64807134	1.67E-15	1.64E-14
ENSBTAG00000018016	*NUPR1*	25	26082047	26083376	2.18407186	5.1458393	1.94E-26	5.74E-25
ENSBTAG00000014122	*FOXRED1*	29	29617419	29625087	2.18756759	6.74409681	1.28E-72	3.39E-69
ENSBTAG00000054765	*PGLYRP4*	3	17126512	17142568	2.21032662	2.84672596	6.61E-49	1.54E-46
ENSBTAG00000011677	*H1-2*	23	31876303	31878150	2.22628994	6.77795024	4.82E-46	8.87E-44
ENSBTAG00000013290	*DYSF*	11	12899573	13123869	2.23294351	7.22533236	1.98E-59	9.37E-57
ENSBTAG00000002635	*PGLYRP1*	18	53513706	53515925	2.24700062	2.17219562	3.56E-16	3.76E-15
ENSBTAG00000018223	*CHI3L1*	16	968434	978542	2.25708517	8.31264111	4.32E-57	1.59E-54
ENSBTAG00000010065	*TRPC5*	X	59975310	60322067	2.27756822	3.910392	3.87E-44	6.10E-42
ENSBTAG00000007169	*P2RX1*	19	24528313	24545951	2.29844203	4.35615943	1.25E-34	7.89E-33
ENSBTAG00000001051	*OSCAR*	18	63283496	63290397	2.32401338	7.21999908	6.67E-45	1.12E-42
ENSBTAG00000013205	*IL1RAP*	1	76577593	76723526	2.33129355	5.61237156	2.63E-47	5.44E-45
ENSBTAG00000006904	*TNS2*	5	26903397	26921895	2.33181233	2.08615516	4.01E-36	2.92E-34
ENSBTAG00000004741	*IL12B*	7	70893041	70912103	2.35487073	2.37995281	3.14E-37	2.63E-35
ENSBTAG00000008389	*HTRA1*	26	42285479	42343309	2.35809337	2.11791885	1.89E-31	8.88E-30
ENSBTAG00000001785	*TGM3*	13	52793538	52836265	2.36208162	9.9423002	1.06E-40	1.23E-38
ENSBTAG00000018446	*GCA*	2	34070353	34087612	2.36293622	3.84814895	1.22E-30	5.35E-29
ENSBTAG00000013201	*ALOX5AP*	12	30108987	30138259	2.40703286	6.9373371	6.31E-47	1.23E-44
ENSBTAG00000020257	*PTPN5*	29	25991833	26050347	2.42337177	4.81324996	5.81E-49	1.37E-46
ENSBTAG00000018134	*AREG*	6	89379645	89391792	2.44626852	2.16046865	3.37E-58	1.31E-55
ENSBTAG00000003920	*TGM1*	10	20902051	20916712	2.4565639	2.68321759	2.95E-36	2.17E-34
ENSBTAG00000010007	*MAPK13*	23	10058481	10067089	2.48414318	4.01377935	1.08E-48	2.47E-46

ENSBTAG00000003519	*NOL3*	18	34807714	34812057	2.48652097	2.56296205	1.27E-59	6.23E-57
ENSBTAG00000005668	*SLC39A8*	6	22459117	22542658	2.49407544	2.92614113	5.96E-46	1.08E-43
ENSBTAG00000012638	*S100A12*	3	17102722	17104173	2.52030644	11.6008924	2.66E-55	8.20E-53
ENSBTAG00000008428	*UPP1*	4	7619639	7648114	2.5492016	7.23254632	9.94E-57	3.47E-54
ENSBTAG00000003353	*SLC6A2*	18	23875374	23939382	2.57574654	2.38872528	1.48E-45	2.58E-43
ENSBTAG00000048737	*DEFB10*	27	6596422	6598413	2.57645919	3.85563287	1.00E-29	4.06E-28
ENSBTAG00000006523	*SOD2*	9	95955338	95966127	2.60697164	9.70477441	7.47E-62	5.21E-59
ENSBTAG00000016566	*ITGA9*	22	10908546	11272063	2.61849499	3.06964927	4.37E-31	1.98E-29
ENSBTAG00000049416	*RAB20*	12	85168908	85200403	2.6539923	5.39992561	3.32E-70	7.33E-67
ENSBTAG00000001292	*LTF*	22	52952571	52986619	2.65838065	3.64223275	9.82E-17	1.10E-15
ENSBTAG00000021887	*DPYS*	14	60148619	60239255	2.73430746	2.12609435	1.13E-30	4.97E-29
ENSBTAG00000019669	*CD163*	5	101786078	101818046	2.74124089	8.22115496	4.50E-62	3.51E-59
ENSBTAG00000046152	*MGAM*	4	105285315	105466927	2.77488972	5.32831653	7.60E-44	1.16E-41
ENSBTAG00000013706	*MEGF9*	8	110032801	110110151	2.81976799	5.63250368	4.75E-60	2.52E-57
ENSBTAG00000017969	*CA4*	19	12803073	12811847	2.87328291	3.69508725	2.72E-69	5.15E-66
ENSBTAG00000015592	*GPR84*	5	25709927	25711851	2.90997466	3.86203237	1.86E-56	6.32E-54
ENSBTAG00000017251	*SLC26A8*	23	9878961	9967961	2.92458937	2.73729509	6.38E-57	2.28E-54
ENSBTAG00000020406	*GPC3*	X	17366350	17829036	2.94499757	3.50945164	1.20E-34	7.61E-33
ENSBTAG00000018280	*SLC28A3*	8	77514711	77587255	2.97496279	6.61909768	4.28E-45	7.27E-43
ENSBTAG00000012640	*S100A8*	3	17085577	17086827	3.01187475	10.0378609	5.43E-58	2.06E-55
ENSBTAG00000020580	*TCN1*	15	83058088	83073458	3.01691809	8.62925931	2.11E-58	8.74E-56
ENSBTAG00000006505	*S100A9*	3	17115128	17117984	3.02955042	11.2625886	1.04E-60	6.26E-58
ENSBTAG00000031950	*RAB3IP*	5	43657005	43713702	3.04020761	6.6902838	1.64E-75	1.09E-71
ENSBTAG00000019330	*PROK2*	22	29859914	29883008	3.05834103	4.70971728	1.28E-62	1.06E-59
ENSBTAG00000002233	*CPNE2*	18	25134033	25166257	3.12972775	4.90756036	1.61E-58	6.88E-56
ENSBTAG00000021240	*DCSTAMP*	14	60265369	60283636	3.13235993	4.52674729	4.67E-47	9.23E-45
ENSBTAG00000006354	*HP*	18	39037402	39043531	3.14928995	9.30133677	3.26E-48	7.32E-46
ENSBTAG00000006221	*ADGRG3*	18	25644891	25674917	3.16049926	5.549692	5.38E-68	7.92E-65
ENSBTAG00000006999	*RYR1*	18	48237459	48365215	3.30921973	5.94180205	1.85E-60	1.02E-57

ENSBTAG00000007239	*TNFAIP6*	2	44747145	44764214	3.34835832	3.88150879	3.97E-55	1.20E-52
ENSBTAG00000020676	*MMP9*	13	74746976	74754303	3.43368711	5.94807756	1.48E-40	1.66E-38
ENSBTAG00000014149	*LCN2*	11	98781893	98785927	3.44623679	6.95188733	8.55E-44	1.29E-41
ENSBTAG00000007901	*ADGRE1*	7	17557794	17633496	3.45389826	8.94007079	4.73E-68	7.83E-65
ENSBTAG00000000377	*BMX*	X	127954722	127992190	3.49095229	5.40899509	8.54E-75	3.77E-71
ENSBTAG00000002996	*SHROOM4*	X	88412754	88541873	3.49107172	3.22509705	2.38E-65	2.63E-62
ENSBTAG00000053557	*DEFB4A*	27	7138873	7140876	3.63970879	3.01214453	2.60E-28	9.14E-27
ENSBTAG00000009773	*KREMEN1*	17	68432890	68476866	3.69612485	6.82415319	8.66E-64	8.20E-61
ENSBTAG00000049808	*IL3RA*	3	119480332	119499353	3.89413651	6.82970855	4.98E-85	6.60E-81
ENSBTAG00000048720		24	61926351	61935647	4.24764936	2.86278283	5.64E-24	1.32E-22
ENSBTAG00000008951	*ALPL*	2	131181416	131245100	4.32698395	7.42489259	1.10E-61	7.29E-59
ENSBTAG00000046158	*CFB*	23	27415355	27421377	4.3602461	5.399988	3.19E-66	3.84E-63
ENSBTAG00000050618		26	42405805	42438497	4.40311183	4.64722978	2.75E-58	1.10E-55
ENSBTAG00000019627	*THY1*	15	29981019	29986913	4.77741378	3.65512439	2.17E-55	6.85E-53
ENSBTAG00000052012		NKLS02000500.1	916	4537	4.78846917	5.79325831	9.54E-37	7.57E-35
ENSBTAG00000010273	*EREG*	6	89306902	89325899	4.81657827	2.21264622	5.84E-62	4.30E-59
ENSBTAG00000054882		24	61844175	61850625	4.98984381	4.3046835	3.10E-38	2.85E-36
ENSBTAG00000051132		24	61907407	61913758	5.1138886	5.24075655	1.08E-42	1.51E-40
ENSBTAG00000039037	*SERPINB4*	NKLS02001094.1	5955	12952	5.20313648	6.36372934	3.02E-51	8.00E-49
ENSBTAG00000048835		24	61862826	61887768	5.31868482	7.65609637	8.80E-48	1.91E-45
ENSBTAG00000049569		NKLS02001931.1	565	6035	5.3721789	6.9730613	1.99E-49	4.88E-47
ENSBTAG00000006343	*IL1R2*	11	6689610	6728985	5.59521591	7.5133198	4.97E-56	1.61E-53
ENSBTAG00000013356	*CATHL3*	22	51579579	51621314	6.08740823	4.37280747	5.49E-33	2.97E-31
ENSBTAG00000031647	*LRG1*	7	19596244	19599321	7.34756702	4.23058149	2.37E-74	7.85E-71

Table 5.3 Ten topmost significantly enriched biological functions associated with differentially expressed genes

Biological function	*P*-value range	Genes involved in the biological function
Inflammatory Response	3.22E-20-2.06E-03	*ADGRE1,ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI, BST1,CA4,CD163,CFB,CHI3L1,DEFB4A/DEFB4B,DPYS,DYSF,EREG,GCA, GPC3,GPR84,HP,HTRA1,IL12B,IL1R2,IL1RAP,IL3RA,ITGA9,LCN2,LRG1,LTF, MAPK13,MGAM,MMP9,MYRF,NRG1,NUPR1,OLFM4,OSCAR,P2RX1, PGLYRP1,PGLYRP4,PROK2,RAB3D,RETN,S100A12,S100A8,S100A9, SERPINB4,SLC39A8,SLC6A2,SOD2,TCN1,TGFA,TGM3,THY1,TNFAIP6, TRPC5,UPP1*
Connective Tissue Disorders	1.55E-14-1.62E-03	*ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CA4,CD163,CFB, CHI3L1,DCSTAMP,DPYS,GCA,GPC3,HP,HTRA1,IL12B,IL1R2,IL3RA,ITGA9, KREMEN1,LCN2,LTF,MMP9,PGLYRP1,PROK2,RETN,S100A12,S100A8, S100A9,SLC39A8,SLC6A2,SOD2,TGFA,TNFAIP6*
Inflammatory Disease	1.55E-14-1.83E-03	*ADGRE1,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CA4,CD163, CFB,CHI3L1,DEFB4A/DEFB4B,DPYS,EREG,GCA, zH1-2,HP,HTRA1,IL12B, IL1R2,IL3RA,ITGA9,LCN2,LRG1,LTF,MGAM,MMP9,NRG1,OLFM4,PGLYRP1, PGLYRP4,PROK2,RETN,S100A12,S100A8,S100A9,SERPINB4,SLC39A8, SLC6A2,SOD2,TCN1,TGFA,TGM3,THY1,TNFAIP6*
Organismal Injury and Abnormalities	1.55E-14-2.07E-03	*ABCA6,ACVR1B,ADGRE1,ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL, ANKRD22,AREG,ARG1,ART5,BMX,BPI,BST1,CA4,CD163,CFB,CHI3L1, CPNE2,DCSTAMP,DEFB4A/DEFB4B,DPYS,DYSF,EREG,FOXRED1,GCA, GLT1D1,GPC3,GPR84,H1-2,HBD,HP,HTRA1,IL12B,IL1R2,IL1RAP,IL3RA, ITGA9,KREMEN1,LCN2,LRG1,LTF,MAPK13,MEGF9,MGAM,MMP9,MN1, MYBPH,MYRF,NOL3,NRG1,NUPR1,OLFM4,OSCAR,P2RX1,PGLYRP1, PGLYRP4,PROK2,PTPN5,RAB20,RAB3D,RAB3IP,RBPMS2,REEP1,RETN, RYR1,S100A12,S100A8,S100A9,SERPINB4,SHROOM4,SLC26A8,SLC28A3, SLC39A8,SLC6A2,SOD2,TCN1,TGFA,TGM1,TGM3,THY1,TNFAIP6,TNS2, TRPC5,UPP1,WIPI1*
Immunological Disease	3.58E-11-2.05E-03	*ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CD163,CFB, CHI3L1,DEFB4A/DEFB4B,GCA,GPC3,GPR84,HP,IL12B,IL1R2,IL3RA,ITGA9, LCN2,LTF,MGAM,MMP9,NRG1,PGLYRP1,PROK2,RETN,S100A12,S100A8, S100A9,SERPINB4,SLC6A2,SOD2,TGFA,TGM3,TNFAIP6*
Infectious Diseases	1.62E-08-1.42E-03	*ALOX5AP,ALPL,BPI,CD163,CFB,DEFB4A/DEFB4B,DYSF,GCA,GPC3,H1-2, HP,IL12B,IL1R2,IL3RA,LCN2,LTF,MGAM,MMP9,MYRF,NRG1,OLFM4,P2RX1, PGLYRP1,RAB3D,RETN,S100A12,S100A8,S100A9,SLC6A2,TCN1*

Respiratory Disease	1.62E-08-1.73E-03	*ABCA6,ACVR1B,ALAS2,ALOX15,ALPL,ANKRD22,AREG,ARG1,BMX,BPI,BST1, CA4,CD163,CFB,CHI3L1,CPNE2,DCSTAMP,DPYS,DYSF,EREG,FOXRED1, GLT1D1,GPC3,GPR84,H1-2,HP,HTRA1,IL12B,IL1RAP,IL3RA,ITGA9,LCN2, LTF,MAPK13,MEGF9,MGAM,MMP9,MN1,MYBPH,MYRF,NOL3,NRG1, NUPR1,OLFM4,PGLYRP1,PGLYRP4,PTPN5,RETN,RYR1,S100A12,S100A8, S100A9,SERPINB4,SHROOM4,SLC6A2,SOD2,TCN1,TGFA,TGM3,THY1, TNFAIP6,TNS2,TRPC5*
Antimicrobial Response	2.01E-08-2.57E-04	*BPI,DEFB4A/DEFB4B,IL12B,LCN2,LTF,PGLYRP1,PGLYRP4,S100A12,S100A8, S100A9*
Psychological Disorders	2.55E-08-1.83E-03	*ALOX15,ARG1,CA4,CFB,CHI3L1,DYSF,HP,HTRA1,IL12B,IL1R2,LCN2,LRG1, LTF,MMP9,NRG1,PTPN5,RYR1,S100A9,SLC6A2,SOD2,TGM1,THY1,UPP1*
Metabolic Disease	1.81E-07-7.5E-04	*ALOX15,ALOX5AP,ALPL,ARG1,BPI,CA4,CFB,CHI3L1,DYSF,GPC3,HBD,HP, HTRA1,IL12B,IL1R2,IL3RA,LCN2,LTF,MGAM,MMP9,PTPN5,RETN,S100A8, S100A9,SLC6A2,SOD2,TGM1,THY1*

Table 5.S1 The significant enriched biological functions associated with differentially expressed genes

Biological function	*P*-value range	Genes involved in the biological function
Inflammatory Response	3.22E-20-2.06E-03	*ADGRE1,ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,BST1, CA4,CD163,CFB,CHI3L1,DEFB4A/DEFB4B,DPYS,DYSF,EREG,GCA,GPC3,GPR84, HP,HTRA1,IL12B,IL1R2,IL1RAP,IL3RA,ITGA9,LCN2,LRG1,LTF,MAPK13,MGAM, MMP9,MYRF,NRG1,NUPR1,OLFM4,OSCAR,P2RX1,PGLYRP1,PGLYRP4,PROK2, RAB3D,RETN,S100A12,S100A8,S100A9,SERPINB4,SLC39A8,SLC6A2,SOD2,TCN1, TGFA,TGM3,THY1,TNFAIP6,TRPC5,UPP1*
Connective Tissue Disorders	1.55E-14-1.62E-03	*ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CA4,CD163,CFB,CHI3L1, DCSTAMP,DPYS,GCA,GPC3,HP,HTRA1,IL12B,IL1R2,IL3RA,ITGA9,KREMEN1, LCN2,LTF,MMP9,PGLYRP1,PROK2,RETN,S100A12,S100A8,S100A9,SLC39A8, SLC6A2,SOD2,TGFA,TNFAIP6*
Inflammatory Disease	1.55E-14-1.83E-03	*ADGRE1,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CA4,CD163,CFB, CHI3L1,DEFB4A/DEFB4B,DPYS,EREG,GCA,H1-2,HP,HTRA1,IL12B,IL1R2,IL3RA,ITGA9,LCN2,LRG1,LTF,MGAM,MMP9,NRG1, OLFM4,PGLYRP1,PGLYRP4,PROK2,RETN,S100A12,S100A8,S100A9,SERPINB4, SLC39A8,SLC6A2,SOD2,TCN1,TGFA,TGM3,THY1,TNFAIP6*
Organismal Injury and Abnormalities	1.55E-14-2.07E-03	*ABCA6,ACVR1B,ADGRE1,ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL,ANKRD22, AREG,ARG1,ART5,BMX,BPI,BST1,CA4,CD163,CFB,CHI3L1,CPNE2,DCSTAMP, DEFB4A/DEFB4B,DPYS,DYSF,EREG,FOXRED1,GCA,GLT1D1,GPC3,GPR84,H1-2, HBD,HP,HTRA1,IL12B,IL1R2,IL1RAP,IL3RA,ITGA9,KREMEN1,LCN2,LRG1,LTF, MAPK13,MEGF9,MGAM,MMP9,MN1,MYBPH,MYRF,NOL3,NRG1,NUPR1,OLFM4, OSCAR,P2RX1,PGLYRP1,PGLYRP4,PROK2,PTPN5,RAB20,RAB3D,RAB3IP, RBPMS2,REEP1,RETN,RYR1,S100A12,S100A8,S100A9,SERPINB4,SHROOM4, SLC26A8,SLC28A3,SLC39A8,SLC6A2,SOD2,TCN1,TGFA,TGM1,TGM3,THY1, TNFAIP6,TNS2,TRPC5,UPP1,WIPI1*
Immunological Disease	3.58E-11-2.05E-03	*ADGRG3,ALAS2,ALOX15,ALOX5AP,ALPL,AREG,ARG1,BMX,BPI,CD163,CFB, CHI3L1,DEFB4A/DEFB4B,GCA,GPC3,GPR84,HP,IL12B,IL1R2,IL3RA,ITGA9, LCN2,LTF,MGAM,MMP9,NRG1,PGLYRP1,PROK2,RETN,S100A12,S100A8,S100A9, SERPINB4,SLC6A2,SOD2,TGFA,TGM3,TNFAIP6*
Infectious Diseases	1.62E-08-1.42E-03	*ALOX5AP,ALPL,BPI,CD163,CFB,DEFB4A/DEFB4B,DYSF,GCA,GPC3,H1-2,HP, IL12B,IL1R2,IL3RA,LCN2,LTF,MGAM,MMP9,MYRF,NRG1,OLFM4,P2RX1, PGLYRP1,RAB3D,RETN,S100A12,S100A8,S100A9,SLC6A2,TCN1*
Respiratory Disease	1.62E-08-1.73E-03	*ABCA6,ACVR1B,ALAS2,ALOX15,ALPL,ANKRD22,AREG,ARG1,BMX,BPI,BST1,CA4, CD163,CFB,CHI3L1,CPNE2,DCSTAMP,DPYS,DYSF,EREG,FOXRED1,GLT1D1, GPC3,GPR84,H1-2,HP,HTRA1,IL12B,IL1RAP,IL3RA,ITGA9,LCN2,LTF,*

		MAPK13,MEGF9,MGAM,MMP9,MN1,MYBPH,MYRF,NOL3,NRG1,NUPR1,OLFM4, PGLYRP1,PGLYRP4,PTPN5,RETN,RYR1,S100A12,S100A8,S100A9,SERPINB4, SHROOM4,SLC6A2,SOD2,TCN1,TGFA,TGM3,THY1,TNFAIP6,TNS2,TRPC5
Antimicrobial Response	2.01E-08-2.57E-04	*BPI,DEFB4A/DEFB4B,IL12B,LCN2,LTF,PGLYRP1,PGLYRP4,S100A12,S100A8, S100A9*
Psychological Disorders	2.55E-08-1.83E-03	*ALOX15,ARG1,CA4,CFB,CHI3L1,DYSF,HP,HTRA1,IL12B,IL1R2,LCN2,LRG1,LTF, MMP9,NRG1,PTPN5,RYR1,S100A9,SLC6A2,SOD2,TGM1,THY1,UPP1*
Metabolic Disease	1.81E-07-7.5E-04	*ALOX15,ALOX5AP,ALPL,ARG1,BPI,CA4,CFB,CHI3L1,DYSF,GPC3,HBD,HP, HTRA1,IL12B,IL1R2,IL3RA,LCN2,LTF,MGAM,MMP9,PTPN5,RETN,S100A8, S100A9,SLC6A2,SOD2,TGM1,THY1*
Endocrine System Disorders	6.75E-06-1.96E-03	*ABCA6,ACVR1B,ADGRE1,ALAS2,ALOX15,ALOX5AP,ALPL,ANKRD22,AREG,ARG1, ART5,BMX,BPI,BST1,CA4,CD163,CFB,CHI3L1,DCSTAMP,DPYS,DYSF,EREG, FOXRED1,GCA,GLT1D1,GPC3,GPR84,H1-2,HBD,HP,HTRA1,IL12B,IL1R2,IL1RAP,IL3RA,ITGA9,KREMEN1,LCN2,LRG1,LTF, MAPK13,MEGF9,MGAM,MMP9,MN1,MYBPH,MYRF,NOL3,NRG1,OLFM4,OSCAR, P2RX1,PGLYRP4,PROK2,PTPN5,RAB20,RAB3D,RBPMS2,REEP1,RETN,RYR1, S100A8,SERPINB4,SHROOM4,SLC26A8,SLC28A3,SLC39A8,SLC6A2,SOD2,TCN1, TGFA,TGM1,TGM3,THY1,TNFAIP6,TNS2,TRPC5,UPP1,WIPI1*
Hereditary Disorder	5.96E-05-1.31E-03	*CA4,CFB,CHI3L1,FOXRED1,GPC3,HP,HTRA1,IL3RA,LCN2,LRG1,MMP9,MN1, MYRF,PROK2,PTPN5,REEP1,RETN,RYR1,S100A12,S100A9,SHROOM4,SLC6A2, SOD2,TGM1*
Nutritional Disease	3.51E-04-1.91E-03	*ADGRE1,AREG,ARG1,CA4,CD163,EREG,IL1RAP,LCN2,LTF,MAPK13,MGAM, MMP9,PGLYRP4,RETN,SLC6A2,TGFA*
Developmental Disorder	7.5E-04-7.5E-04	*LCN2,MMP9*
Hematological Disease	7.75E-04-7.75E-04	*CFB,IL12B,LCN2,LTF,MMP9,P2RX1,S100A9*

Table 5.S2 Differentially expressed genes between BRD and non-BRD animals in feedlot 1

Gene ID	Gene name	Chr	Start (bp)	End (bp)	logFC	logCPM	*P*-Value	FDR
ENSBTAG00000054844	*HBA1*	25	219512	220318	-4.2973366	10.2568414	1.83E-23	7.71E-22
ENSBTAG00000051412	*HBA*	25	216448	217264	-4.2783593	10.2297791	2.10E-23	8.76E-22
ENSBTAG00000013178	*ALAS2*	X	92485200	92512358	-3.6103259	4.82602073	7.74E-41	3.58E-38
ENSBTAG00000037644	*HBB*	15	48362236	48363996	-3.3646968	10.3183578	3.77E-32	4.94E-30
ENSBTAG00000011990	*ALOX15*	19	26697231	26705809	-3.2522801	4.67377917	1.30E-23	5.55E-22
ENSBTAG00000048626		19	14399906	14401309	-2.5987282	4.3164336	2.76E-29	2.35E-27
ENSBTAG00000001032	*PYGM*	29	42985611	42997675	-2.2929225	3.33328741	7.89E-39	2.74E-36
ENSBTAG00000014269	*SCARB1*	17	50929375	51022457	2.0451812	7.69285595	1.56E-47	1.44E-44
ENSBTAG00000008648	*PRDX5*	29	42613257	42616366	2.06314739	7.46569807	1.83E-28	1.42E-26
ENSBTAG00000019428	*CCR1*	22	53225174	53231223	2.07442803	8.76031592	2.50E-24	1.16E-22
ENSBTAG00000045854		18	62462232	62476548	2.0772604	9.88027762	3.13E-30	3.06E-28
ENSBTAG00000006685	*IL10*	16	4550836	4555318	2.09595882	2.61777573	2.34E-23	9.69E-22
ENSBTAG00000016204	*C1RL*	5	103273388	103298036	2.09935865	6.58260581	1.99E-38	6.50E-36
ENSBTAG00000015032	*CD14*	7	51762895	51765768	2.11674682	7.55592213	1.59E-27	1.09E-25
ENSBTAG00000019567	*IL21R*	25	24941958	25002153	2.14437884	7.19622647	8.38E-48	9.69E-45
ENSBTAG00000011037	*RBPMS2*	10	45313259	45339402	2.2919649	3.05337159	1.10E-21	3.93E-20
ENSBTAG00000035081	*SERINC2*	2	122233162	122259028	2.30961334	7.63510967	5.44E-30	5.13E-28
ENSBTAG00000021647	*FCAR*	18	62527485	62537893	2.31287175	5.42434663	1.03E-23	4.42E-22
ENSBTAG00000018517	*VLDLR*	8	41826623	41860757	2.34188981	5.07370421	6.19E-22	2.27E-20
ENSBTAG00000051383	*DEFB7*	27	6676076	6678281	2.35681888	4.56083724	2.22E-18	5.17E-17
ENSBTAG00000001511	*BCL6*	1	79558314	79611183	2.36728084	8.62532537	1.77E-38	5.99E-36
ENSBTAG00000010328	*KCNK17*	23	13166550	13179213	2.37664071	4.71405008	9.64E-52	1.91E-48
ENSBTAG00000050072		11	98685584	98686780	2.43401674	2.49742416	2.61E-14	3.65E-13
ENSBTAG00000003944	*CATIP*	2	106378600	106384257	2.44485054	2.22172121	3.88E-37	1.02E-34
ENSBTAG00000010047	*TIAM2*	9	91723944	91866359	2.49004614	4.65145371	4.43E-26	2.62E-24
ENSBTAG00000006608	*GGT5*	17	71443740	71453369	2.49391529	5.98245285	1.07E-28	8.55E-27
ENSBTAG00000006156	*BST1*	6	110872406	110901795	2.49479041	5.71011438	4.65E-30	4.48E-28
ENSBTAG00000008124	*GK*	X	112861218	112935556	2.49939178	6.02675446	1.89E-42	1.01E-39
ENSBTAG00000020892	*IL2RA*	13	17455360	17500714	2.50649074	6.93085166	5.53E-50	7.67E-47
ENSBTAG00000010464	*MN1*	17	67173546	67221633	2.51938721	2.6135965	2.01E-38	6.50E-36
ENSBTAG00000013290	*DYSF*	11	12899573	13123869	2.52995181	7.51607068	1.96E-33	3.13E-31
ENSBTAG00000008851		18	57488548	57494764	2.57648382	6.82309135	9.10E-32	1.10E-29
ENSBTAG00000013555	*ACVR1B*	5	27884739	27915481	2.58162706	6.15735641	1.21E-34	2.29E-32

ENSBTAG00000052465		9	42992439	42996483	2.59490249	7.16693166	2.16E-24	1.02E-22
ENSBTAG00000023648	*ART5*	15	51386896	51443955	2.5973771	7.40605672	3.63E-37	9.70E-35
ENSBTAG00000014122	*FOXRED1*	29	29617419	29625087	2.5984207	7.09967922	1.20E-47	1.27E-44
ENSBTAG00000002148	*RAB3D*	7	15736270	15747432	2.59990187	7.56604139	3.05E-35	6.42E-33
ENSBTAG00000018016	*NUPR1*	25	26082047	26083376	2.60415958	5.54503137	2.07E-19	5.58E-18
ENSBTAG00000033748	*IL18RAP*	11	7203304	7233720	2.60556797	5.48416939	1.39E-33	2.32E-31
ENSBTAG00000020430	*GLT1D1*	17	48170769	48286681	2.60569099	3.21120558	2.19E-25	1.18E-23
ENSBTAG00000037826		15	51474978	51495512	2.65750801	3.05319112	1.61E-31	1.86E-29
ENSBTAG00000014069	*PDK4*	4	12881889	12895362	2.65775179	5.92019635	5.71E-35	1.17E-32
ENSBTAG00000030424	*CLEC1A*	5	99860312	99880739	2.67070157	2.61775197	1.94E-13	2.46E-12
ENSBTAG00000016477	*PEAK3*	7	21274513	21280458	2.67657098	5.11716636	2.80E-36	6.95E-34
ENSBTAG00000054765	*PGLYRP4*	3	17126512	17142568	2.68395549	3.19950095	1.04E-37	3.13E-35
ENSBTAG00000013368	*ANKRD22*	26	10506966	10534508	2.69776247	4.22625509	1.52E-22	6.00E-21
ENSBTAG00000004150	*NRG1*	27	28529266	28651519	2.7230818	5.07904789	5.11E-31	5.63E-29
ENSBTAG00000010065	*TRPC5*	X	59975310	60322067	2.7287576	4.00345186	2.63E-25	1.39E-23
ENSBTAG00000011465	*MYBPH*	16	957169	965458	2.73364827	6.4809191	5.58E-38	1.72E-35
ENSBTAG00000006921	*ABCA6*	19	61421551	61482846	2.73990641	6.84004875	1.28E-47	1.27E-44
ENSBTAG00000007169	*P2RX1*	19	24528313	24545951	2.76542813	4.88197075	6.91E-19	1.73E-17
ENSBTAG00000045808	*IL17REL*	5	119366047	119390141	2.83812653	2.62546178	5.86E-17	1.14E-15
ENSBTAG00000022779	*OLFM4*	12	10641503	10665121	2.84013624	3.70949773	1.96E-21	6.76E-20
ENSBTAG00000018134	*AREG*	6	89379645	89391792	2.86258141	2.49231926	5.05E-31	5.60E-29
ENSBTAG00000006904	*TNS2*	5	26903397	26921895	2.88678435	2.50111355	1.98E-21	6.79E-20
ENSBTAG00000000783	*TGFA*	11	13861144	13975868	2.90139822	3.04966321	5.06E-37	1.30E-34
ENSBTAG00000011677	*H1-2*	23	31876303	31878150	2.92811547	7.00726852	6.28E-35	1.26E-32
ENSBTAG00000013201	*ALOX5AP*	12	30108987	30138259	2.95881911	7.35645499	2.47E-28	1.90E-26
ENSBTAG00000020257	*PTPN5*	29	25991833	26050347	2.97487725	5.11780993	4.55E-32	5.79E-30
ENSBTAG00000018223	*CHI3L1*	16	968434	978542	2.99720233	8.79119794	3.12E-55	2.17E-51
ENSBTAG00000013205	*IL1RAP*	1	76577593	76723526	3.01042415	5.94716852	6.55E-27	4.23E-25
ENSBTAG00000001785	*TGM3*	13	52793538	52836265	3.02278062	10.3869376	1.01E-34	1.94E-32
ENSBTAG00000014046	*BPI*	13	67172370	67201998	3.03626779	6.30989356	4.59E-12	4.83E-11
ENSBTAG00000003920	*TGM1*	10	20902051	20916712	3.06804087	3.15341894	1.26E-20	3.88E-19
ENSBTAG00000002635	*PGLYRP1*	18	53513706	53515925	3.0807196	2.88253583	4.33E-10	3.50E-09
ENSBTAG00000018446	*GCA*	2	34070353	34087612	3.09043557	4.33977129	3.44E-21	1.14E-19
ENSBTAG00000003519	*NOL3*	18	34807714	34812057	3.09100227	2.99551523	2.66E-41	1.32E-38
ENSBTAG00000001051	*OSCAR*	18	63283496	63290397	3.11383381	7.70693664	2.12E-43	1.22E-40
ENSBTAG00000003353	*SLC6A2*	18	23875374	23939382	3.16452408	2.84511352	6.55E-43	3.64E-40

ENSBTAG00000004716	*RETN*	7	16513147	16514562	3.18032812	5.74769745	1.15E-13	1.51E-12
ENSBTAG00000008428	*UPP1*	4	7619639	7648114	3.18317751	7.60620298	1.43E-39	5.52E-37
ENSBTAG00000010007	*MAPK13*	23	10058481	10067089	3.21213355	4.46914928	1.09E-27	7.71E-26
ENSBTAG00000006523	*SOD2*	9	95955338	95966127	3.21903451	10.1428464	1.85E-50	2.85E-47
ENSBTAG00000012638	*S100A12*	3	17102722	17104173	3.23871478	11.9809943	9.81E-34	1.72E-31
ENSBTAG00000013706	*MEGF9*	8	110032801	110110151	3.25291167	5.8873535	4.70E-30	4.50E-28
ENSBTAG00000008389	*HTRA1*	26	42285479	42343309	3.30062971	2.44259414	9.05E-24	3.90E-22
ENSBTAG00000048737	*DEFB10*	27	6596422	6598413	3.31426917	4.41138357	5.70E-20	1.61E-18
ENSBTAG00000005668	*SLC39A8*	6	22459117	22542658	3.32110747	3.30072448	1.04E-33	1.79E-31
ENSBTAG00000049416	*RAB20*	12	85168908	85200403	3.33713107	5.66321856	3.24E-52	7.49E-49
ENSBTAG00000017969	*CA4*	19	12803073	12811847	3.41311077	3.97922501	2.85E-55	2.17E-51
ENSBTAG00000004741	*IL12B*	7	70893041	70912103	3.41683604	2.92274451	2.53E-32	3.48E-30
ENSBTAG00000019669	*CD163*	5	101786078	101818046	3.47992853	8.66021642	8.81E-36	2.11E-33
ENSBTAG00000015592	*GPR84*	5	25709927	25711851	3.49054978	4.39291106	1.01E-33	1.75E-31
ENSBTAG00000017251	*SLC26A8*	23	9878961	9967961	3.49176436	3.21189608	2.55E-29	2.18E-27
ENSBTAG00000020406	*GPC3*	X	17366350	17829036	3.49788532	3.62061423	3.42E-23	1.40E-21
ENSBTAG00000032068	*PLA2G4F*	10	37533181	37547704	3.49864846	2.39761523	2.55E-39	9.51E-37
ENSBTAG00000001292	*LTF*	22	52952571	52986619	3.56864745	4.48067791	8.23E-11	7.30E-10
ENSBTAG00000002233	*CPNE2*	18	25134033	25166257	3.61212146	5.33770111	9.28E-30	8.48E-28
ENSBTAG00000016566	*ITGA9*	22	10908546	11272063	3.65826711	3.71930323	1.23E-23	5.27E-22
ENSBTAG00000031950	*RAB3IP*	5	43657005	43713702	3.68072539	7.04207079	5.46E-47	4.21E-44
ENSBTAG00000019330	*PROK2*	22	29859914	29883008	3.70684804	4.94377801	1.87E-35	4.25E-33
ENSBTAG00000020580	*TCN1*	15	83058088	83073458	3.74727764	9.13666359	2.61E-38	8.24E-36
ENSBTAG00000046152	*MGAM*	4	105285315	105466927	3.77057077	5.7578948	8.12E-29	6.60E-27
ENSBTAG00000038532	*FOLR3*	15	51782270	51788157	3.77102735	2.42521719	2.33E-11	2.24E-10
ENSBTAG00000006505	*S100A9*	3	17115128	17117984	3.78303888	11.7813729	2.08E-35	4.58E-33
ENSBTAG00000018280	*SLC28A3*	8	77514711	77587255	3.81736019	7.12643211	9.01E-32	1.10E-29
ENSBTAG00000006354	*HP*	18	39037402	39043531	3.82895489	9.8732448	1.15E-25	6.44E-24
ENSBTAG00000012640	*S100A8*	3	17085577	17086827	3.85988315	10.4517459	5.49E-32	6.80E-30
ENSBTAG00000002996	*SHROOM4*	X	88412754	88541873	4.03659463	3.53651657	3.15E-37	8.56E-35
ENSBTAG00000006221	*ADGRG3*	18	25644891	25674917	4.0495256	6.02515919	1.75E-50	2.85E-47
ENSBTAG00000007239	*TNFAIP6*	2	44747145	44764214	4.12434796	4.20592922	6.65E-31	7.16E-29
ENSBTAG00000007901	*ADGRE1*	7	17557794	17633496	4.13527429	9.52174038	1.24E-37	3.66E-35
ENSBTAG00000021240	*DCSTAMP*	14	60265369	60283636	4.14693808	5.14457708	4.18E-35	8.65E-33
ENSBTAG00000014149	*LCN2*	11	98781893	98785927	4.17123564	7.75525651	5.58E-24	2.48E-22
ENSBTAG00000006999	*RYR1*	18	48237459	48365215	4.18280317	6.56993497	5.98E-39	2.13E-36

ENSBTAG00000020676	*MMP9*	13	74746976	74754303	4.25226299	6.56583352	1.04E-21	3.75E-20
ENSBTAG00000049808	*IL3RA*	3	119480332	119499353	4.27203036	7.38165778	5.77E-53	2.00E-49
ENSBTAG00000000377	*BMX*	X	127954722	127992190	4.35406757	5.72334627	2.01E-47	1.75E-44
ENSBTAG00000009773	*KREMEN1*	17	68432890	68476866	4.41817659	7.41941193	1.33E-37	3.85E-35
ENSBTAG00000053557	*DEFB4A*	27	7138873	7140876	4.72411959	3.94443641	1.40E-16	2.62E-15
ENSBTAG00000046158	*CFB*	23	27415355	27421377	5.07295915	6.13366715	5.20E-45	3.14E-42
ENSBTAG00000050618		26	42405805	42438497	5.08634489	5.32979311	1.58E-34	2.93E-32
ENSBTAG00000008951	*ALPL*	2	131181416	131245100	5.26645087	7.91570538	9.70E-33	1.43E-30
ENSBTAG00000019627	*THY1*	15	29981019	29986913	5.59422236	4.31342278	5.00E-30	4.75E-28
ENSBTAG00000048720		24	61926351	61935647	6.04526541	3.61944597	3.67E-21	1.21E-19
ENSBTAG00000010273	*EREG*	6	89306902	89325899	6.05111123	2.76189694	1.94E-35	4.33E-33
ENSBTAG00000039037	*SERPINB4*	NKLS02001094.1	5955	12952	6.56022101	7.00535753	2.99E-35	6.38E-33
ENSBTAG00000052012		NKLS02000500.1	916	4537	6.59024775	6.49339707	3.12E-31	3.52E-29
ENSBTAG00000054882	*IL1R2*	24	61844175	61850625	6.66547067	5.03213692	4.49E-28	3.32E-26
ENSBTAG00000006343		11	6689610	6728985	6.78390898	8.40125835	6.68E-32	8.20E-30
ENSBTAG00000013356	*CATHL3*	22	51579579	51621314	6.79789647	5.51283742	9.53E-16	1.60E-14
ENSBTAG00000051132		24	61907407	61913758	6.80917139	5.93934354	1.19E-31	1.42E-29
ENSBTAG00000048835		24	61862826	61887768	6.86172982	8.30800721	4.29E-34	7.73E-32
ENSBTAG00000049569		NKLS02001931.1	565	6035	6.89066714	7.67022948	2.98E-34	5.45E-32
ENSBTAG00000002976	*CD177*	18	51546288	51557282	7.06281866	4.43321029	2.37E-15	3.78E-14
ENSBTAG00000024852	*CATHL2*	22	51641063	51642757	8.00271765	5.96569378	2.29E-14	3.22E-13
ENSBTAG00000031647	*LRG1*	7	19596244	19599321	8.33047218	5.11724925	4.98E-46	3.46E-43

Table 5.S3 Differentially expressed genes between BRD and non-BRD animals in feedlot 2

Gene ID	Gene name	Chr	Start (bp)	End (bp)	logFC	logCPM	*P*-Value	FDR
ENSBTAG00000037644	*HBB*	15	48362236	48363996	-3.3313062	12.6174759	3.36E-05	0.00081402
ENSBTAG00000051412	*HBA*	25	216448	217264	-3.2274209	12.5576244	0.00041715	0.00618737
ENSBTAG00000054844	*HBA1*	25	219512	220318	-3.2271536	12.5442137	0.00041478	0.00618464
ENSBTAG00000006683	*ZFYVE28*	6	116265797	116364222	-3.0506406	4.16002225	1.88E-13	6.43E-11
ENSBTAG00000013178	*ALAS2*	X	92485200	92512358	-2.773232	6.55951193	0.00010051	0.00199821
ENSBTAG00000008945	*SDSL*	17	61005835	61018993	-2.429167	2.4812053	1.71E-10	2.39E-08
ENSBTAG00000048817	*SSX5*	X	86550779	86560943	-2.2953572	3.03922109	0.00013688	0.0025311
ENSBTAG00000023666		1	42936757	42941412	-2.2446964	2.17864132	2.67E-05	0.00066925
ENSBTAG00000010464	*MN1*	17	67173546	67221633	2.04100252	2.57788149	5.80E-13	1.66E-10
ENSBTAG00000022779	*OLFM4*	12	10641503	10665121	2.04305947	2.81370769	2.03E-07	1.23E-05
ENSBTAG00000033748	*IL18RAP*	11	7203304	7233720	2.04431193	5.02817798	9.89E-10	1.14E-07
ENSBTAG00000002635	*PGLYRP1*	18	53513706	53515925	2.05649417	2.07530631	1.90E-06	7.76E-05
ENSBTAG00000013205	*IL1RAP*	1	76577593	76723526	2.07593902	5.48048561	3.07E-10	4.02E-08
ENSBTAG00000021240	*DCSTAMP*	14	60265369	60283636	2.08905388	3.52450437	3.98E-07	2.13E-05
ENSBTAG00000012638	*S100A12*	3	17102722	17104173	2.10931851	11.5346472	5.92E-13	1.66E-10
ENSBTAG00000003519	*NOL3*	18	34807714	34812057	2.12067719	2.26424501	1.66E-13	5.95E-11
ENSBTAG00000013201	*ALOX5AP*	12	30108987	30138259	2.13487295	6.74574051	6.08E-11	9.69E-09
ENSBTAG00000007169	*P2RX1*	19	24528313	24545951	2.14209392	4.0406586	4.19E-13	1.25E-10
ENSBTAG00000018446	*GCA*	2	34070353	34087612	2.1449927	3.5814541	3.16E-06	0.00011553
ENSBTAG00000011037	*RBPMS2*	10	45313259	45339402	2.17621438	3.04954011	7.08E-11	1.12E-08
ENSBTAG00000006999	*RYR1*	18	48237459	48365215	2.22541545	5.09029834	2.23E-14	1.01E-11
ENSBTAG00000013290	*DYSF*	11	12899573	13123869	2.2359714	7.12160127	7.78E-13	2.11E-10
ENSBTAG00000018134	*AREG*	6	89379645	89391792	2.25624408	2.04085164	5.04E-16	3.72E-13
ENSBTAG00000008389	*HTRA1*	26	42285479	42343309	2.26295351	2.05763963	1.24E-10	1.81E-08
ENSBTAG00000020257	*PTPN5*	29	25991833	26050347	2.26382172	4.73365293	7.39E-11	1.15E-08
ENSBTAG00000049416	*RAB20*	12	85168908	85200403	2.26561588	5.15606209	2.49E-11	4.72E-09
ENSBTAG00000008428	*UPP1*	4	7619639	7648114	2.29100809	6.97550656	4.92E-11	8.21E-09
ENSBTAG00000019669	*CD163*	5	101786078	101818046	2.34176127	8.09449315	2.49E-19	8.74E-16
ENSBTAG00000010007	*MAPK13*	23	10058481	10067089	2.38598678	3.94875145	2.24E-22	1.05E-18
ENSBTAG00000018280	*SLC28A3*	8	77514711	77587255	2.39104332	5.9741236	4.16E-07	2.19E-05
ENSBTAG00000046152	*MGAM*	4	105285315	105466927	2.4190759	5.11070344	2.42E-13	7.89E-11
ENSBTAG00000020580	*TCN1*	15	83058088	83073458	2.4442418	8.21479133	4.78E-10	5.93E-08
ENSBTAG00000014149	*LCN2*	11	98781893	98785927	2.44567427	5.92534488	1.72E-17	2.42E-14

ENSBTAG00000012640	*S100A8*	3	17085577	17086827	2.53368911	9.97236459	9.57E-17	1.03E-13
ENSBTAG00000015592	*GPR84*	5	25709927	25711851	2.55001628	3.27906263	1.13E-12	2.87E-10
ENSBTAG00000031950	*RAB3IP*	5	43657005	43713702	2.5720683	6.35104255	1.94E-12	4.45E-10
ENSBTAG00000048737	*DEFB10*	27	6596422	6598413	2.58624064	3.73008651	1.18E-11	2.36E-09
ENSBTAG00000006505	*S100A9*	3	17115128	17117984	2.59480996	10.9989727	1.42E-15	9.84E-13
ENSBTAG00000003920	*TGM1*	10	20902051	20916712	2.69735948	2.53300975	1.78E-10	2.47E-08
ENSBTAG00000017969	*CA4*	19	12803073	12811847	2.71539183	3.47709747	2.99E-10	3.95E-08
ENSBTAG00000013706	*MEGF9*	8	110032801	110110151	2.76130293	5.4912256	2.46E-12	5.47E-10
ENSBTAG00000030542	*KRT25*	19	40945832	40953601	2.78521268	2.20961455	3.07E-06	0.00011295
ENSBTAG00000006221	*ADGRG3*	18	25644891	25674917	2.81196772	5.21290453	2.51E-14	1.07E-11
ENSBTAG00000007901	*ADGRE1*	7	17557794	17633496	2.87149171	8.46415973	1.51E-15	9.84E-13
ENSBTAG00000006354	*HP*	18	39037402	39043531	2.91144279	9.18915673	1.91E-15	1.12E-12
ENSBTAG00000020676	*MMP9*	13	74746976	74754303	2.95666303	5.70616247	1.71E-16	1.60E-13
ENSBTAG00000017251	*SLC26A8*	23	9878961	9967961	2.97349195	2.57032764	1.54E-15	9.84E-13
ENSBTAG00000009773	*KREMEN1*	17	68432890	68476866	3.05295603	6.30284439	4.55E-13	1.33E-10
ENSBTAG00000019330	*PROK2*	22	29859914	29883008	3.118044	4.8970174	4.29E-16	3.34E-13
ENSBTAG00000000377	*BMX*	X	127954722	127992190	3.13915197	5.36563341	6.94E-15	3.48E-12
ENSBTAG00000002233	*CPNE2*	18	25134033	25166257	3.23703352	4.87264695	1.98E-16	1.63E-13
ENSBTAG00000049808	*IL3RA*	3	119480332	119499353	3.28565671	6.29977977	6.15E-18	1.08E-14
ENSBTAG00000002996	*SHROOM4*	X	88412754	88541873	3.30069481	3.17782575	2.38E-11	4.64E-09
ENSBTAG00000007239	*TNFAIP6*	2	44747145	44764214	3.37583136	4.03797458	2.37E-15	1.28E-12
ENSBTAG00000038532	*FOLR3*	15	51782270	51788157	3.56655729	2.31178514	3.54E-07	1.93E-05
ENSBTAG00000046158	*CFB*	23	27415355	27421377	3.88357977	4.73882412	3.20E-19	8.97E-16
ENSBTAG00000048720		24	61926351	61935647	3.90684514	2.6226416	4.10E-07	2.16E-05
ENSBTAG00000008951	*ALPL*	2	131181416	131245100	4.0775995	7.48684535	5.87E-15	3.05E-12
ENSBTAG00000050618		26	42405805	42438497	4.14181549	3.93242761	8.66E-19	1.97E-15
ENSBTAG00000019627	*THY1*	15	29981019	29986913	4.34843311	3.65005702	1.68E-15	1.03E-12
ENSBTAG00000052012		NKLS02000500.1	916	4537	4.35423906	5.52250449	1.17E-09	1.30E-07
ENSBTAG00000013356	*CATHL3*	22	51579579	51621314	4.49790509	2.09755897	1.62E-12	3.78E-10
ENSBTAG00000054882		24	61844175	61850625	4.55978418	4.12153344	2.48E-11	4.72E-09
ENSBTAG00000051132		24	61907407	61913758	4.69519508	5.14328519	1.10E-12	2.87E-10
ENSBTAG00000006343	*IL1R2*	11	6689610	6728985	4.72909304	6.54130568	8.75E-17	1.02E-13
ENSBTAG00000039037	*SERPINB4*	NKLS02001094.1	5955	12952	4.91245874	6.29279909	9.87E-14	3.64E-11

ENSBTAG00000049569		NKLS02001931.1	565	6035	4.9156393	6.82459895	2.39E-13	7.89E-11
ENSBTAG00000048835		24	61862826	61887768	4.96763452	7.68066012	2.42E-14	1.06E-11
ENSBTAG00000031647	*LRG1*	7	19596244	19599321	5.69037075	2.77832152	9.85E-19	1.97E-15

Table 5.S4 Differentially expressed genes between BRD and non-BRD animals in feedlot 3

Gene ID	Gene name	Chr	Start (bp)	End (bp)	logFC	logCPM	*P*-Value	FDR
ENSBTAG00000037644	*HBB*	15	48362236	48363996	-2.4610853	10.9910341	5.58E-06	9.43E-05
ENSBTAG00000054844	*HBA1*	25	219512	220318	-2.3558354	10.8789008	0.00054774	0.00391512
ENSBTAG00000051412	*HBA*	25	216448	217264	-2.3408418	10.8748729	0.00058089	0.00410557
ENSBTAG00000013103	*COL1A1*	19	36457658	36474513	-2.2497554	2.10531141	0.00058759	0.00413532
ENSBTAG00000003353	*SLC6A2*	18	23875374	23939382	2.04508943	2.02940069	2.25E-11	2.11E-09
ENSBTAG00000019669	*CD163*	5	101786078	101818046	2.05711042	7.68609912	1.71E-22	2.02E-19
ENSBTAG00000050072		11	98685584	98686780	2.05951156	2.51460641	9.78E-09	4.22E-07
ENSBTAG00000020257	*PTPN5*	29	25991833	26050347	2.06473288	4.56528543	7.72E-14	1.36E-11
ENSBTAG00000011677	*H1-2*	23	31876303	31878150	2.07702475	6.73942837	3.98E-12	4.32E-10
ENSBTAG00000046152	*MGAM*	4	105285315	105466927	2.1177211	4.91260054	1.57E-13	2.54E-11
ENSBTAG00000010065	*TRPC5*	X	59975310	60322067	2.16389242	3.96116393	3.54E-16	1.07E-13
ENSBTAG00000006354	*HP*	18	39037402	39043531	2.19786697	8.49216483	1.47E-14	2.98E-12
ENSBTAG00000017251	*SLC26A8*	23	9878961	9967961	2.27587554	2.20497117	8.33E-20	4.74E-17
ENSBTAG00000006505	*S100A9*	3	17115128	17117984	2.29066818	10.7417117	3.14E-17	1.12E-14
ENSBTAG00000014149	*LCN2*	11	98781893	98785927	2.31824499	5.92209667	1.10E-12	1.43E-10
ENSBTAG00000049416	*RAB20*	12	85168908	85200403	2.3207807	5.31700656	5.12E-21	3.83E-18
ENSBTAG00000006523	*SOD2*	9	95955338	95966127	2.3372362	9.4506076	1.43E-16	4.64E-14
ENSBTAG00000021240	*DCSTAMP*	14	60265369	60283636	2.34805146	3.95801971	6.20E-13	8.73E-11
ENSBTAG00000015592	*GPR84*	5	25709927	25711851	2.36792851	3.44807493	1.13E-15	2.82E-13
ENSBTAG00000005668	*SLC39A8*	6	22459117	22542658	2.38895254	2.84803044	7.07E-14	1.26E-11
ENSBTAG00000017969	*CA4*	19	12803073	12811847	2.3936347	3.49069294	2.35E-17	8.58E-15
ENSBTAG00000012640	*S100A8*	3	17085577	17086827	2.41285684	9.67528266	1.02E-17	3.93E-15
ENSBTAG00000006221	*ADGRG3*	18	25644891	25674917	2.43186869	5.03641916	2.24E-19	1.10E-16
ENSBTAG00000018280	*SLC28A3*	8	77514711	77587255	2.44998412	6.25229155	3.07E-15	7.39E-13
ENSBTAG00000020580	*TCN1*	15	83058088	83073458	2.47982961	8.20013624	1.19E-17	4.46E-15
ENSBTAG00000020676	*MMP9*	13	74746976	74754303	2.56952049	5.24356952	1.74E-10	1.27E-08
ENSBTAG00000006999	*RYR1*	18	48237459	48365215	2.62857618	5.35516813	1.42E-16	4.64E-14

ENSBTAG00000002233	*CPNE2*	18	25134033	25166257	2.64968344	4.46085806	1.48E-21	1.32E-18
ENSBTAG00000019330	*PROK2*	22	29859914	29883008	2.65674983	4.50428397	4.04E-19	1.92E-16
ENSBTAG00000013706	*MEGF9*	8	110032801	110110151	2.69749793	5.52061047	8.37E-26	1.70E-22
ENSBTAG00000031950	*RAB3IP*	5	43657005	43713702	2.71787798	6.56882955	2.10E-25	3.73E-22
ENSBTAG00000007901	*ADGRE1*	7	17557794	17633496	2.83449419	8.31689294	3.32E-25	5.25E-22
ENSBTAG00000020406	*GPC3*	X	17366350	17829036	2.88606854	4.00591981	2.63E-09	1.33E-07
ENSBTAG00000007239	*TNFAIP6*	2	44747145	44764214	3.00019776	3.67215149	1.23E-18	5.29E-16
ENSBTAG00000009773	*KREMEN1*	17	68432890	68476866	3.13059082	6.27154975	1.44E-20	8.53E-18
ENSBTAG00000000377	*BMX*	X	127954722	127992190	3.15872392	5.18735718	2.34E-26	1.11E-22
ENSBTAG00000002996	*SHROOM4*	X	88412754	88541873	3.21720041	2.98543432	6.38E-21	4.32E-18
ENSBTAG00000046158	*CFB*	23	27415355	27421377	3.31808654	4.61711846	6.68E-16	1.90E-13
ENSBTAG00000008951	*ALPL*	2	131181416	131245100	3.64066428	6.86734475	7.70E-26	1.70E-22
ENSBTAG00000049808	*IL3RA*	3	119480332	119499353	3.64623916	6.39580944	6.14E-26	1.70E-22
ENSBTAG00000050618		26	42405805	42438497	3.67340999	3.78500568	7.15E-19	3.28E-16
ENSBTAG00000019627	*THY1*	15	29981019	29986913	3.68128695	2.6580191	2.14E-14	4.22E-12
ENSBTAG00000006343	*IL1R2*	11	6689610	6728985	3.89683235	6.09712733	6.64E-15	1.45E-12
ENSBTAG00000052012		NKLS02000500.1	916	4537	3.98339891	4.67448468	4.67E-12	4.96E-10
ENSBTAG00000054882		24	61844175	61850625	4.10213514	3.12521762	2.62E-13	3.97E-11
ENSBTAG00000051132		24	61907407	61913758	4.31078223	4.11077112	1.45E-15	3.55E-13
ENSBTAG00000039037	*SERPINB4*	NKLS02001094.1	5955	12952	4.63915587	5.4509929	2.14E-19	1.09E-16
ENSBTAG00000049569		NKLS02001931.1	565	6035	4.65415732	5.91411498	8.72E-19	3.88E-16
ENSBTAG00000048835		24	61862826	61887768	4.65476445	6.60846966	7.04E-18	2.78E-15
ENSBTAG00000031647	*LRG1*	7	19596244	19599321	5.83537733	2.83595109	9.81E-20	5.37E-17

Table 5.S5 Differentially expressed genes between BRD and non-BRD animals in feedlot 4

Gene ID	Gene name	Chr	Start (bp)	End (bp)	logFC	logCPM	*P*-Value	FDR
ENSBTAG00000030542	*KRT25*	19	40945832	40953601	-3.1631419	3.19472604	6.07E-06	0.00028546
ENSBTAG00000011990	*ALOX15*	19	26697231	26705809	-2.8434332	5.18852128	4.02E-08	5.51E-06
ENSBTAG00000048626		19	14399906	14401309	-2.4309794	4.6724847	8.01E-08	9.67E-06
ENSBTAG00000000273	*IL5RA*	22	23199161	23239003	-2.4064042	3.36953694	1.63E-06	0.00010059
ENSBTAG00000020406	*GPC3*	X	17366350	17829036	2.04982109	2.84286597	1.09E-06	7.67E-05
ENSBTAG00000018016	*NUPR1*	25	26082047	26083376	2.07109327	4.90114664	0.00012781	0.00300328
ENSBTAG00000011037	*RBPMS2*	10	45313259	45339402	2.09675603	2.75170321	6.95E-11	3.09E-08
ENSBTAG00000005359	*TGFB2*	16	21808003	21900539	2.17843158	3.31002987	0.00013051	0.00304657
ENSBTAG00000009773	*KREMEN1*	17	68432890	68476866	2.18682267	5.58969684	1.08E-11	6.69E-09
ENSBTAG00000007901	*ADGRE1*	7	17557794	17633496	2.19899112	7.93373867	1.56E-12	1.31E-09
ENSBTAG00000000377	*BMX*	X	127954722	127992190	2.21441204	4.67720178	6.60E-11	3.03E-08
ENSBTAG00000017969	*CA4*	19	12803073	12811847	2.24418964	3.30553407	6.62E-10	1.89E-07
ENSBTAG00000039037	*SERPINB4*	NKLS0200 1094.1	5955	12952	2.2733472	4.60775457	0.00046975	0.00790682
ENSBTAG00000046158	*CFB*	23	27415355	27421377	2.33122503	3.30552444	1.32E-13	1.57E-10
ENSBTAG00000002996	*SHROOM4*	X	88412754	88541873	2.54630902	2.5834404	4.70E-11	2.39E-08
ENSBTAG00000008951	*ALPL*	2	131181416	131245100	2.57744764	6.05698002	7.59E-09	1.33E-06
ENSBTAG00000006343	*IL1R2*	11	6689610	6728985	2.82220547	5.17177951	5.38E-11	2.64E-08
ENSBTAG00000049808	*IL3RA*	3	119480332	119499353	2.97607174	5.65420191	1.06E-20	1.52E-16
ENSBTAG00000050618		26	42438497		3.58513307	26.624644	2.23E-05	
ENSBTAG00000004757	*LTBP4*	18	49897503	49924389	4.27249106	2.81632176	2.49E-10	8.65E-08
ENSBTAG00000031647	*LRG1*	7	19596244	19599321	5.98911774	2.68838879	6.40E-14	9.33E-11

Table 5.S6 Common and unique differential expressed genes between four feedlots

Feedlot	Number of overlapping genes	Gene ID
FEEDLOT 1, FEEDLOT 2, FEEDLOT 3, FEEDLOT 4	12	ENSBTAG00000046158, ENSBTAG00000049808, ENSBTAG00000006343, ENSBTAG00000002996, ENSBTAG00000007901, ENSBTAG00000039037, ENSBTAG00000008951, ENSBTAG00000009773, ENSBTAG00000000377, ENSBTAG00000050618, ENSBTAG00000031647, ENSBTAG00000017969
FEEDLOT 1, FEEDLOT 2, FEEDLOT 4	1	ENSBTAG00000020406
FEEDLOT 1, FEEDLOT 2, FEEDLOT 3	30	ENSBTAG00000020676, ENSBTAG00000006221, ENSBTAG00000037644, ENSBTAG00000046152, ENSBTAG00000014149, ENSBTAG00000006354, ENSBTAG00000019330, ENSBTAG00000007239, ENSBTAG00000049569, ENSBTAG00000051412, ENSBTAG00000012640, ENSBTAG00000006505, ENSBTAG00000015592, ENSBTAG00000052012, ENSBTAG00000017251, ENSBTAG00000018280, ENSBTAG00000013706, ENSBTAG00000002233, ENSBTAG00000048835, ENSBTAG00000054844, ENSBTAG00000020580, ENSBTAG00000021240, ENSBTAG00000019627, ENSBTAG00000054882, ENSBTAG00000031950, ENSBTAG00000049416, ENSBTAG00000051132, ENSBTAG00000006999, ENSBTAG00000020257, ENSBTAG00000019669
FEEDLOT 1, FEEDLOT 3, FEEDLOT 4	1	ENSBTAG00000011037
FEEDLOT 1, FEEDLOT 2	6	ENSBTAG00000011677, ENSBTAG00000010065, ENSBTAG00000050072, ENSBTAG00000005668, ENSBTAG00000003353, ENSBTAG00000006523
FEEDLOT 1, FEEDLOT 4	3	ENSBTAG00000048626, ENSBTAG00000011990, ENSBTAG00000018016
FEEDLOT 3, FEEDLOT 4	1	ENSBTAG00000030542
FEEDLOT 1, FEEDLOT 3	21	ENSBTAG00000008389, ENSBTAG00000033748, ENSBTAG00000003519, ENSBTAG00000008428, ENSBTAG00000038532, ENSBTAG00000048720, ENSBTAG00000018134, ENSBTAG00000007169, ENSBTAG00000010464, ENSBTAG00000013290, ENSBTAG00000013201, ENSBTAG00000010007, ENSBTAG00000018446, ENSBTAG00000013178, ENSBTAG00000012638, ENSBTAG00000022779, ENSBTAG00000003920, ENSBTAG00000013205, ENSBTAG00000013356, ENSBTAG00000002635, ENSBTAG00000048737
FEEDLOT 2	1	ENSBTAG00000013103
FEEDLOT 4	3	ENSBTAG00000005359, ENSBTAG00000004757, ENSBTAG00000000273
FEEDLOT 1	53	ENSBTAG00000020430, ENSBTAG00000030424, ENSBTAG00000010273, ENSBTAG00000053557, ENSBTAG00000004741, ENSBTAG00000001292,

	ENSBTAG00000016477, ENSBTAG00000002148, ENSBTAG00000052465, ENSBTAG00000006608, ENSBTAG00000015032, ENSBTAG00000023648, ENSBTAG00000014069, ENSBTAG00000021647, ENSBTAG00000020892, ENSBTAG00000010328, ENSBTAG00000008851, ENSBTAG00000013555, ENSBTAG00000006156, ENSBTAG00000019567, ENSBTAG00000024852, ENSBTAG00000014046, ENSBTAG00000037826, ENSBTAG00000002976, ENSBTAG00000004716, ENSBTAG00000004150, ENSBTAG00000000783, ENSBTAG00000008648, ENSBTAG00000016204, ENSBTAG00000010047, ENSBTAG00000006685, ENSBTAG00000051383, ENSBTAG00000001785, ENSBTAG00000014269, ENSBTAG00000014122, ENSBTAG00000013368, ENSBTAG00000001032, ENSBTAG00000001511, ENSBTAG00000018223, ENSBTAG00000035081, ENSBTAG00000018517, ENSBTAG00000008124, ENSBTAG00000045854, ENSBTAG00000011465, ENSBTAG00000019428, ENSBTAG00000032068, ENSBTAG00000054765, ENSBTAG00000006904, ENSBTAG00000006921, ENSBTAG00000003944, ENSBTAG00000001051, ENSBTAG00000016566, ENSBTAG00000045808

Table 5.S7 Common and unique differential expressed genes between combined population and separated population

Population	Number of overlapping genes
Combined Separated	96
Separated	40
Combined	5

Table 5.S8 cis-eQTLs and associated differentially expressed genes

SNP	Gene ID	b	*P*-value	FDR
rs136463215	ENSBTAG00000050072	-1.151441192	1.25E-07	0.00361096
rs132778261	ENSBTAG00000050072	-1.036920098	6.63E-07	0.00593164
BovineNovelSNP6258	ENSBTAG00000023648	-0.936823393	9.45E-07	0.00593164
rs136117999	ENSBTAG00000050072	-0.959500463	1.31E-06	0.00593164
rs135114400	ENSBTAG00000050072	-1.015796131	1.72E-06	0.00593164
rs135509473	ENSBTAG00000050072	-1.015796131	1.72E-06	0.00593164
rs135903006	ENSBTAG00000050072	-1.015796131	1.72E-06	0.00593164
rs43710072	ENSBTAG00000050072	-0.993328162	1.79E-06	0.00593164
rs137181196	ENSBTAG00000050072	-1.004675437	2.04E-06	0.00593164
rs209419196	ENSBTAG00000014046	-1.004018411	2.06E-06	0.00593164
rs109814352	ENSBTAG00000050072	-0.984223847	3.84E-06	0.00882835
rs136875779	ENSBTAG00000050072	-0.987365313	3.98E-06	0.00882835
rs136639428	ENSBTAG00000050072	-0.987365313	3.98E-06	0.00882835
rs379381979	ENSBTAG00000050072	-1.020528922	4.79E-06	0.00986986
rs378845799	ENSBTAG00000037826	1.496328968	6.53E-06	0.01169291
rs134318440	ENSBTAG00000050072	-1.011799556	7.58E-06	0.01169291
rs110027202	ENSBTAG00000007169	1.333801699	7.91E-06	0.01169291
BovineNovelSNP7867	ENSBTAG00000001051	1.046877853	8.38E-06	0.01169291
BPI-1	ENSBTAG00000014046	-0.96921043	9.19E-06	0.01169291
BovineNovelSNP6247	ENSBTAG00000037826	1.356986622	9.43E-06	0.01169291
rs134992936	ENSBTAG00000050072	-0.863103168	9.71E-06	0.01169291
rs135887479	ENSBTAG00000050072	-0.863103168	9.71E-06	0.01169291
rs380015858	ENSBTAG00000037826	1.257427145	1.08E-05	0.01169291
rs41764149	ENSBTAG00000023648	-0.652011472	1.13E-05	0.01169291
rs109866590	ENSBTAG00000007169	1.345077876	1.18E-05	0.01169291
BovineNovelSNP6248	ENSBTAG00000037826	1.342154121	1.25E-05	0.01169291
rs480991415	ENSBTAG00000007169	1.102918492	1.44E-05	0.01169291
rs211272757	ENSBTAG00000007169	1.273035574	1.55E-05	0.01169291
BovineNovelSNP6257	ENSBTAG00000023648	-0.707169856	1.67E-05	0.01169291
rs109253689	ENSBTAG00000050072	-0.810288529	1.69E-05	0.01169291
BovineHD2700002052	ENSBTAG00000053557	0.979333349	1.91E-05	0.01169291
rs133689837	ENSBTAG00000050072	-0.757932887	2.00E-05	0.01169291
BovineNovelSNP5007	ENSBTAG00000050072	1.105340547	2.08E-05	0.01169291
rs133288538	ENSBTAG00000050072	-0.833040473	2.12E-05	0.01169291
rs378845799	ENSBTAG00000023648	1.470652322	2.12E-05	0.01169291
rs384834845	ENSBTAG00000037826	1.227015394	2.15E-05	0.01169291
rs133584270	ENSBTAG00000006523	1.1179606	2.17E-05	0.01169291
BovineHD2700002049	ENSBTAG00000053557	0.966348033	2.18E-05	0.01169291
rs381615356	ENSBTAG00000007169	1.262189968	2.29E-05	0.01169291
rs133559965	ENSBTAG00000050072	-0.816126161	2.40E-05	0.01169291
rs110649799	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs384236795	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs109675415	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs211337252	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs208096245	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs210027361	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs384661513	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs378701182	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291

rs108992565	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs109123599	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs383570468	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs379189766	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs380912239	ENSBTAG00000037826	1.299574946	2.47E-05	0.01169291
rs41764156	ENSBTAG00000023648	-0.626272323	2.58E-05	0.01169291
rs41764148	ENSBTAG00000023648	-0.622027786	2.67E-05	0.01169291
rs109469511	ENSBTAG00000007169	1.250065307	2.72E-05	0.01169291
rs134218499	ENSBTAG00000050072	-0.861241626	2.75E-05	0.01169291
rs109963601	ENSBTAG00000007169	1.242414453	3.15E-05	0.01169291
BovineNovelSNP1874	ENSBTAG00000015592	1.675763745	3.31E-05	0.01169291
rs111005070	ENSBTAG00000007169	1.336629559	3.44E-05	0.01169291
rs110221216	ENSBTAG00000007169	1.336629559	3.44E-05	0.01169291
BovineNovelSNP6236	ENSBTAG00000037826	1.054219123	3.62E-05	0.01169291
rs478517851	ENSBTAG00000007169	1.381533337	3.72E-05	0.01169291
rs135920660	ENSBTAG00000050072	-0.929104675	3.77E-05	0.01169291
rs109423917	ENSBTAG00000037826	1.344214818	3.82E-05	0.01169291
rs109494290	ENSBTAG00000007169	1.25652957	3.85E-05	0.01169291
rs209697538	ENSBTAG00000007169	1.322189214	3.93E-05	0.01169291
rs208918103	ENSBTAG00000007169	1.33360881	3.99E-05	0.01169291
rs379204276	ENSBTAG00000007169	1.343404088	4.14E-05	0.01169291
rs208597045	ENSBTAG00000007169	1.343404088	4.14E-05	0.01169291
rs208392503	ENSBTAG00000007169	1.321680177	4.20E-05	0.01169291
BovineNovelSNP2890	ENSBTAG00000006156	0.912743673	4.27E-05	0.01169291
BovineHD1100028701	ENSBTAG00000050072	0.813649541	4.36E-05	0.01169291
rs134253779	ENSBTAG00000013368	0.567168129	4.49E-05	0.01169291
rs210855945	ENSBTAG00000037826	1.264310659	4.62E-05	0.01169291
BovineHD1900007133	ENSBTAG00000007169	0.692888452	4.86E-05	0.01169291
BovineHD4100000445	ENSBTAG00000013205	1.407531059	5.19E-05	0.01169291
rs380015858	ENSBTAG00000023648	1.2064919	5.21E-05	0.01169291
rs109720430	ENSBTAG00000007169	1.261237356	5.26E-05	0.01169291
rs109276468	ENSBTAG00000007169	1.261237356	5.26E-05	0.01169291
rs41764146	ENSBTAG00000023648	-0.594293896	5.30E-05	0.01169291
rs384378514	ENSBTAG00000050072	-1.109339172	5.63E-05	0.01169291
Hapmap24097-BTA-155074	ENSBTAG00000054882	1.416580892	5.73E-05	0.01169291
rs110435405	ENSBTAG00000007169	1.252064338	5.88E-05	0.01169291
rs207625164	ENSBTAG00000037826	1.142558449	5.99E-05	0.01169291
rs209077424	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs211382737	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208590366	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110355995	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209396627	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208158771	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109894072	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209873016	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs211364835	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109657179	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109951644	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs380500408	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109799358	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291

rs454128726	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
BovineNovelSNP8305	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs474090479	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs385386703	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209382483	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs210664159	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs385534586	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110183067	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110470192	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs380870463	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110216795	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209640330	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109395581	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208397929	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109407915	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109713305	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209824393	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208937080	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs207552699	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs384304611	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110319551	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110521456	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109295626	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs207981923	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs381459677	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs382981281	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209377890	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs211438878	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs380673294	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs379586566	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110358901	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209560186	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs210852240	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208617602	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110768706	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109931725	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs435667495	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs717603562	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs455545137	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs380133354	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs386036038	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs382260499	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs381517349	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs210696845	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs109806956	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs208243367	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs209805262	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs110062927	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs210976463	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs210313551	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291
rs461463504	ENSBTAG00000007169	1.276717138	6.04E-05	0.01169291

Hapmap24097-BTA-155074	ENSBTAG00000048720	1.321538763	6.16E-05	0.01184515
BovineHD1100028669	ENSBTAG00000050072	0.798580366	6.62E-05	0.01250736
rs382687054	ENSBTAG00000007169	1.179416396	6.76E-05	0.01250736
rs210550945	ENSBTAG00000007169	1.266277044	6.98E-05	0.01250736
rs41764160	ENSBTAG00000023648	-0.711786121	7.05E-05	0.01250736
rs110984867	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs109097777	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs210292124	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs109639120	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs108987961	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs378032218	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs209609982	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs210896530	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs211352602	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs109023860	ENSBTAG00000007169	1.274903424	7.11E-05	0.01250736
rs470959062	ENSBTAG00000037826	1.213000822	7.22E-05	0.01259037
rs137449581	ENSBTAG00000003519	0.763691372	7.25E-05	0.01259037
rs210528754	ENSBTAG00000007169	1.173498571	7.33E-05	0.01259037
rs209557475	ENSBTAG00000007169	1.173498571	7.33E-05	0.01259037
rs136573913	ENSBTAG00000050072	-1.005072876	7.55E-05	0.0128134
rs133943504	ENSBTAG00000050072	-1.005072876	7.55E-05	0.0128134
BovineNovelSNP6235	ENSBTAG00000037826	1.053600465	7.63E-05	0.01287224
rs41764164	ENSBTAG00000023648	0.61885078	7.96E-05	0.01335045
rs109083521	ENSBTAG00000007169	1.291300239	8.17E-05	0.01346885
rs110992329	ENSBTAG00000007169	1.291300239	8.17E-05	0.01346885
rs109362135	ENSBTAG00000007169	1.249591053	8.17E-05	0.01346885
rs132727016	ENSBTAG00000050072	-0.743411407	8.38E-05	0.01373781
rs209788846	ENSBTAG00000007169	0.952819635	8.52E-05	0.0138117
rs109802603	ENSBTAG00000007169	1.166456092	8.59E-05	0.0138117
BovineNovelSNP6247	ENSBTAG00000023648	1.258223876	8.60E-05	0.0138117
BovineHD2400018165	ENSBTAG00000054882	1.518050232	8.61E-05	0.0138117
rs134880342	ENSBTAG00000023648	-0.694835939	8.83E-05	0.01407982
BovineHD1300019256	ENSBTAG00000014046	-0.81598874	9.33E-05	0.01479579
rs137796850	ENSBTAG00000006523	1.053592941	9.41E-05	0.01482432
BovineHD1500015132	ENSBTAG00000023648	0.649413975	9.45E-05	0.01482432
rs208819931	ENSBTAG00000007169	1.241528314	9.57E-05	0.0149297
rs109022863	ENSBTAG00000007169	1.281082597	9.76E-05	0.0150106
rs110486937	ENSBTAG00000007169	1.281082597	9.76E-05	0.0150106
rs208453447	ENSBTAG00000007169	1.15938873	9.99E-05	0.0150106
rs379837335	ENSBTAG00000037826	1.226069828	0.000103075	0.0150106
rs383741409	ENSBTAG00000023648	0.632542159	0.000104347	0.0150106
BovineNovelSNP8060	ENSBTAG00000001051	0.981606217	0.000107214	0.0150106
rs211049632	ENSBTAG00000050072	-1.174543696	0.000108578	0.0150106
rs110727928	ENSBTAG00000007169	1.172329021	0.000111995	0.0150106
BovineNovelSNP6248	ENSBTAG00000023648	1.241393737	0.000112037	0.0150106
ARS-BFGL-NGS-113310	ENSBTAG00000050072	-0.99972776	0.000112422	0.0150106
rs109839719	ENSBTAG00000037826	1.271824674	0.000113028	0.0150106
BovineNovelSNP8059	ENSBTAG00000001051	1.047264388	0.000113443	0.0150106
rs109523587	ENSBTAG00000007169	1.167866187	0.000115592	0.0150106

rs109243850	ENSBTAG00000037826	1.239199826	0.000116126	0.0150106
rs378119501	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs209454814	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110484719	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110291359	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110590171	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs379579470	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs386005521	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs17871932	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs384757898	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs382814052	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110655477	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs381009841	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs384751641	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs109163269	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs111030273	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs381355412	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs109301790	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110002677	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs380747717	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs384642823	ENSBTAG00000037826	1.105883787	0.00011634	0.0150106
rs110609944	ENSBTAG00000003519	0.677118868	0.000118396	0.0150106
rs384834845	ENSBTAG00000023648	1.160398966	0.000119075	0.0150106
rs378940497	ENSBTAG00000037826	1.198108907	0.000120871	0.0150106
rs110649799	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs384236795	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs109675415	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs211337252	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs208096245	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs210027361	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs384661513	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs378701182	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs108992565	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs109123599	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs383570468	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs379189766	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs380912239	ENSBTAG00000023648	1.235484462	0.000122232	0.0150106
rs110492696	ENSBTAG00000037826	1.038771405	0.000122922	0.01503143
rs108947221	ENSBTAG00000007169	1.254015856	0.000129994	0.01582124
rs110691164	ENSBTAG00000003519	0.751742806	0.000132517	0.01582124
BovineNovelSNP6242	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
rs43125505	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
rs43125490	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
BovineNovelSNP6244	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
BovineNovelSNP6245	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
BovineNovelSNP6246	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
BovineNovelSNP6253	ENSBTAG00000037826	1.168370704	0.000134315	0.01582124
rs17871916	ENSBTAG00000003920	0.820068048	0.000138339	0.01622897
BovineHD1500015154	ENSBTAG00000023648	-0.624430743	0.000140768	0.01638074
ARS-BFGL-NGS-73444	ENSBTAG00000023648	-0.624430743	0.000140768	0.01638074

BovineNovelSNP6238	ENSBTAG00000023648	0.671736621	0.000147079	0.01704634
rs110498821	ENSBTAG00000007169	1.135310679	0.000157707	0.01820507
rs379128552	ENSBTAG00000037826	1.036323909	0.000159825	0.01827918
BovineHD2400018165	ENSBTAG00000051132	1.492279332	0.000160174	0.01827918
rs109198864	ENSBTAG00000007169	1.256584751	0.000160249	0.01827918
rs133366258	ENSBTAG00000053557	-1.140421309	0.000161869	0.01839126
rs379388109	ENSBTAG00000037826	1.196331857	0.000165851	0.01870033
rs210855945	ENSBTAG00000023648	1.218365684	0.000165885	0.01870033
BovineNovelSNP8061	ENSBTAG00000001051	1.020655115	0.000167209	0.01877618
rs441131550	ENSBTAG00000046158	-1.669538827	0.000173629	0.01934652
rs378772585	ENSBTAG00000046158	-1.669538827	0.000173629	0.01934652
BovineNovelSNP3300	ENSBTAG00000031647	-2.057585956	0.000181864	0.02018617
BovineNovelSNP2644	ENSBTAG00000005668	-0.952859279	0.000185021	0.02031634
rs41765522	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
rs41765523	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
rs41765524	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
rs41766436	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
rs41766444	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
rs109012205	ENSBTAG00000023648	-0.679671285	0.000187964	0.02031634
BovineNovelSNP8151	ENSBTAG00000001051	0.913417547	0.000189995	0.02045917
rs526797318	ENSBTAG00000007169	1.11123967	0.000195585	0.02082796
rs382497480	ENSBTAG00000007169	1.11123967	0.000195585	0.02082796
rs110389111	ENSBTAG00000007169	1.11123967	0.000195585	0.02082796
rs137152628	ENSBTAG00000050072	-0.871778226	0.000201168	0.02130784
rs211589616	ENSBTAG00000006523	1.005311061	0.000201568	0.02130784
rs109276195	ENSBTAG00000037826	1.165894655	0.000202337	0.02131114
BovineNovelSNP7882	ENSBTAG00000001051	0.953002591	0.00020716	0.02173971
BovineHD1100028832	ENSBTAG00000050072	-0.703476907	0.00021111	0.02207401
ARS-BFGL-NGS-5022	ENSBTAG00000023648	0.591897661	0.000214272	0.02232376
rs209051289	ENSBTAG00000007169	1.100294845	0.000220507	0.02280863
rs209727009	ENSBTAG00000007169	1.100294845	0.000220507	0.02280863
BovineHD2400018203	ENSBTAG00000048720	1.387928169	0.000226113	0.02314209
rs211035768	ENSBTAG00000007169	1.165544261	0.000226499	0.02314209
BovineHD1500014907	ENSBTAG00000023648	0.602004472	0.000227043	0.02314209
ARS-BFGL-NGS-59463	ENSBTAG00000023648	0.602004472	0.000227043	0.02314209
rs135908577	ENSBTAG00000050072	-0.925734864	0.00022774	0.02314209
rs441961624	ENSBTAG00000052465	-0.986343555	0.000230942	0.02338506
rs135413666	ENSBTAG00000050072	-0.736443565	0.000237453	0.02396038
rs109423917	ENSBTAG00000023648	1.252706917	0.000238549	0.02398702
rs209436475	ENSBTAG00000007169	1.227996191	0.000244387	0.02448873
BovineHD1500015115	ENSBTAG00000023648	0.635477851	0.00024702	0.02458257
rs110262321	ENSBTAG00000003519	0.691405243	0.000247027	0.02458257
rs135236972	ENSBTAG00000015592	0.716104131	0.000248064	0.02460093
BovineNovelSNP3258	ENSBTAG00000007901	-1.313844662	0.000249904	0.02469853
rs135908199	ENSBTAG00000006523	1.107655927	0.00025542	0.02505791
rs110955377	ENSBTAG00000007169	1.187972519	0.000256145	0.02505791
rs110088639	ENSBTAG00000007169	1.187972519	0.000256145	0.02505791
rs210582705	ENSBTAG00000006523	0.980956898	0.000259068	0.02517316
rs211353086	ENSBTAG00000006523	0.980956898	0.000259068	0.02517316
rs109906421	ENSBTAG00000006523	-0.702352509	0.000274169	0.02650275

BovineNovelSNP7827	ENSBTAG00000001051	0.991459031	0.000274588	0.02650275
rs450088264	ENSBTAG00000007169	1.180731623	0.000278132	0.02675535
rs207625164	ENSBTAG00000023648	1.077809325	0.000285897	0.02741098
rs385165514	ENSBTAG00000006523	-0.699941884	0.00028918	0.02756553
rs211483835	ENSBTAG00000007169	1.042513079	0.000289419	0.02756553
BovineNovelSNP8064	ENSBTAG00000001051	1.07776921	0.000290383	0.02756629
Hapmap24097-BTA-155074	ENSBTAG00000051132	1.308673575	0.000292258	0.02765334
rs384814481	ENSBTAG00000007169	1.173508083	0.000297142	0.02802361
rs210895989	ENSBTAG00000007169	1.129308562	0.000300122	0.02821246
rs110660745	ENSBTAG00000007169	1.171900291	0.000301512	0.02825111
rs136568050	ENSBTAG00000050072	-0.759583298	0.000302688	0.02826951
BovineHD2400018101	ENSBTAG00000048720	-1.309083629	0.000306477	0.02853102
rs381050941	ENSBTAG00000049808	0.961622851	0.000316496	0.02934307
rs110203907	ENSBTAG00000007169	1.144703893	0.000317233	0.02934307
rs381006828	ENSBTAG00000046158	-1.618569151	0.000321168	0.02961207
rs209822745	ENSBTAG00000007169	1.168312594	0.000322433	0.02963407
rs17871713	ENSBTAG00000037826	1.105699567	0.00032539	0.0298109
rs384240836	ENSBTAG00000037826	1.047266865	0.000328364	0.02998812
BTB-00604291	ENSBTAG00000023648	0.597033046	0.000336727	0.03026428
BovineHD1500014999	ENSBTAG00000023648	-0.597033046	0.000336727	0.03026428
rs41768418	ENSBTAG00000023648	-0.597033046	0.000336727	0.03026428
BovineHD1500015059	ENSBTAG00000023648	-0.597033046	0.000336727	0.03026428
rs109413789	ENSBTAG00000037826	0.989095261	0.000338645	0.03026428
rs378112010	ENSBTAG00000006523	0.832600051	0.000338888	0.03026428
rs210900104	ENSBTAG00000007169	0.955737377	0.000340826	0.03026428
rs109792791	ENSBTAG00000007169	0.955737377	0.000340826	0.03026428
rs209772652	ENSBTAG00000007169	0.955737377	0.000340826	0.03026428
rs379368071	ENSBTAG00000007169	1.1682631	0.000351858	0.03105278
rs210642963	ENSBTAG00000007169	1.1682631	0.000351858	0.03105278
rs384545049	ENSBTAG00000007169	0.958186262	0.000354189	0.03114051
rs137604908	ENSBTAG00000046152	-1.33989126	0.00035501	0.03114051
rs135282255	ENSBTAG00000050072	-0.702948787	0.000361932	0.03151878
rs110684121	ENSBTAG00000007169	1.164108491	0.000362577	0.03151878
rs378940497	ENSBTAG00000023648	1.157157577	0.000365289	0.03151878
rs207689213	ENSBTAG00000007169	1.167366273	0.000371589	0.03151878
rs209253469	ENSBTAG00000007169	1.167366273	0.000371589	0.03151878
rs210292339	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
rs380443848	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
rs383490727	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
rs210197442	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
rs208212613	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
rs211490355	ENSBTAG00000007169	0.668506357	0.000372668	0.03151878
BovineHD2400018165	ENSBTAG00000048835	1.421417041	0.000373101	0.03151878
rs208957111	ENSBTAG00000007169	1.074017241	0.000375312	0.03151878
rs380327479	ENSBTAG00000046158	-2.017963814	0.000375344	0.03151878
rs136938332	ENSBTAG00000013368	0.516289627	0.000375705	0.03151878
BovineNovelSNP3270	ENSBTAG00000007901	-1.177962003	0.00038276	0.03198242
rs208015502	ENSBTAG00000007169	1.150029735	0.000386772	0.03198242
rs110473764	ENSBTAG00000007169	1.150029735	0.000386772	0.03198242
rs211602540	ENSBTAG00000007169	1.150029735	0.000386772	0.03198242

rs209059328	ENSBTAG00000007169	1.150029735	0.000386772	0.03198242
rs43125499	ENSBTAG00000023648	-0.642431741	0.000392423	0.03235697
rs210214722	ENSBTAG00000007169	1.148303068	0.000402295	0.03307647
BPI-2	ENSBTAG00000014046	-0.799893603	0.000427395	0.03504036
rs209886375	ENSBTAG00000049808	2.018315435	0.000434097	0.03544752
rs41764145	ENSBTAG00000023648	-0.601132133	0.000434818	0.03544752
rs210926622	ENSBTAG00000007169	1.141048947	0.000437896	0.03559784
rs110492696	ENSBTAG00000023648	0.990181993	0.000445686	0.03602509
BovineNovelSNP8063	ENSBTAG00000001051	1.071118484	0.000445859	0.03602509
rs382462514	ENSBTAG00000037826	0.966862432	0.000446896	0.03602509
BovineNovelSNP3270	ENSBTAG00000004716	-0.749992196	0.0004505	0.03621068
rs137238434	ENSBTAG00000006523	1.045044909	0.000458873	0.03621068
rs109839719	ENSBTAG00000023648	1.201125777	0.00047261	0.03621068
rs17871713	ENSBTAG00000023648	1.116762509	0.0004767	0.03621068
BovineHD1300019252	ENSBTAG00000014046	0.714746869	0.000477925	0.03621068
ARS-BFGL-BAC-13210	ENSBTAG00000014046	-0.714746869	0.000477925	0.03621068
rs134805892	ENSBTAG00000050072	-0.942123265	0.000479238	0.03621068
rs110956138	ENSBTAG00000037826	1.030997003	0.000480583	0.03621068
rs378119501	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs209454814	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs110484719	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs110291359	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs110590171	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs379579470	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs386005521	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs17871932	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs384757898	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs382814052	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs110655477	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs381009841	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs384751641	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs109163269	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs111030273	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs381355412	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs109301790	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs110002677	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs380747717	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs384642823	ENSBTAG00000023648	1.044347834	0.000484332	0.03621068
rs379544660	ENSBTAG00000037826	1.049901381	0.000498184	0.03713197
rs132835009	ENSBTAG00000015592	0.736168496	0.000499227	0.03713197
BovineNovelSNP3240	ENSBTAG00000004716	-0.738490866	0.000516562	0.03832255
rs109852396	ENSBTAG00000037826	1.089684946	0.000518017	0.03833191
rs135757859	ENSBTAG00000006523	0.902703564	0.000532709	0.03901846
BovineNovelSNP8062	ENSBTAG00000001051	1.055324975	0.00053402	0.03901846
BovineNovelSNP6646	ENSBTAG00000010464	-0.713461954	0.000534025	0.03901846
rs109370010	ENSBTAG00000007169	1.155335935	0.000535407	0.03901846
rs110050587	ENSBTAG00000007169	1.155335935	0.000535407	0.03901846
rs110697789	ENSBTAG00000007169	1.155335935	0.000535407	0.03901846
rs443132523	ENSBTAG00000007169	1.158704354	0.00054021	0.03926935
rs470959062	ENSBTAG00000023648	1.104105201	0.000547717	0.03971501

rs211664177	ENSBTAG00000007169	1.025559079	0.000551447	0.03988524
rs41764149	ENSBTAG00000037826	-0.500333125	0.000564785	0.0407478
rs385877417	ENSBTAG00000046158	-1.957115928	0.000568585	0.04091968
rs109110987	ENSBTAG00000037826	1.086567953	0.0005714	0.04102001
BovineHD0200038361	ENSBTAG00000008951	1.095270955	0.000588135	0.04211662
BovineHD1800010610	ENSBTAG00000003519	-0.662402622	0.000599608	0.04278606
rs379388109	ENSBTAG00000023648	1.136086191	0.000602061	0.04278606
BovineHD0200012931	ENSBTAG00000007239	-0.782030421	0.000603296	0.04278606
rs379837335	ENSBTAG00000023648	1.130917531	0.000603414	0.04278606
rs137193464	ENSBTAG00000050072	-0.622590402	0.000613639	0.04340444
rs208520654	ENSBTAG00000007169	0.996150427	0.000617591	0.04345865
BovineNovelSNP7611	ENSBTAG00000001051	1.07550156	0.000618525	0.04345865
rs110776111	ENSBTAG00000006523	-0.643019641	0.000618923	0.04345865
rs476887482	ENSBTAG00000009773	-1.56897358	0.000620678	0.04347611
BTA-118330-no-rs	ENSBTAG00000017969	0.723577835	0.000649187	0.04536294
rs134790017	ENSBTAG00000013368	0.488364676	0.000663854	0.04627576
rs134045305	ENSBTAG00000006990	0.995327553	0.000677703	0.04712727
rs385450627	ENSBTAG00000037826	0.998914885	0.000692034	0.04800822
rs385226415	ENSBTAG00000037826	1.05241891	0.000697833	0.04829437
rs384404605	ENSBTAG00000046158	-1.544651071	0.000710332	0.04882173
rs481915313	ENSBTAG00000046158	-1.924734957	0.000710413	0.04882173
rs382268066	ENSBTAG00000046158	-1.405140848	0.000710528	0.04882173

Table 5.S9 trans-eQTLs and associated differentially expressed genes

SNP	Gene ID	b	*P*-value	FDR
rs207554348	ENSBTAG00000000783	-1.567618021	2.49E-09	0.00074946
rs207554348	ENSBTAG00000001051	-1.679171656	9.76E-10	0.00033515
rs207554348	ENSBTAG00000001292	-1.745873419	2.36E-09	0.00072162
rs472302580	ENSBTAG00000001292	1.971267192	5.24E-08	0.01037913
rs207554348	ENSBTAG00000001785	-1.848084539	1.28E-10	6.47E-05
rs207554348	ENSBTAG00000002148	-1.642985365	5.19E-14	9.98E-07
BovineNovelSNP579	ENSBTAG00000002148	1.101158625	9.54E-08	0.01730064
rs207554348	ENSBTAG00000002233	-2.096358163	6.97E-11	3.83E-05
BovineNovelSNP579	ENSBTAG00000002233	1.7420936	2.58E-09	0.00076183
rs207554348	ENSBTAG00000002635	-1.378935733	1.24E-08	0.00293961
rs472302580	ENSBTAG00000002635	1.810910269	6.85E-10	0.00025832
rs207554348	ENSBTAG00000003353	-1.800046315	1.05E-09	0.00035305
BovineNovelSNP2890	ENSBTAG00000003353	1.373225367	1.57E-07	0.02551795
rs207554348	ENSBTAG00000003519	-1.681341522	1.33E-10	6.57E-05
BovineNovelSNP579	ENSBTAG00000003519	1.332073305	2.71E-08	0.00584919
rs207554348	ENSBTAG00000003920	-1.80490207	1.28E-12	1.97E-06
BovineNovelSNP579	ENSBTAG00000003920	1.306988544	3.46E-08	0.00723057
rs207554348	ENSBTAG00000004150	-1.520876975	3.77E-10	0.00015763
rs207554348	ENSBTAG00000004716	-1.409380216	4.53E-10	0.00018546
rs472302580	ENSBTAG00000004716	1.474145913	1.89E-07	0.02906562
rs207554348	ENSBTAG00000004741	-1.700008009	7.36E-10	0.00026702
BovineNovelSNP579	ENSBTAG00000004741	1.313696014	2.08E-07	0.03171393
rs207554348	ENSBTAG00000005668	-1.97014086	2.32E-13	1.11E-06
ARS-BFGL-NGS-80280	ENSBTAG00000005668	-1.615340993	3.12E-07	0.04378099
rs207554348	ENSBTAG00000006156	-1.624978444	2.22E-11	1.33E-05
BovineNovelSNP579	ENSBTAG00000006156	1.173886567	1.89E-07	0.02906562
rs207554348	ENSBTAG00000006221	-2.193621609	8.55E-11	4.44E-05
BovineNovelSNP579	ENSBTAG00000006221	1.702167154	4.14E-08	0.00846375
rs207554348	ENSBTAG00000006343	-3.367262729	9.12E-12	7.41E-06
rs207554348	ENSBTAG00000006354	-2.227279685	9.25E-12	7.41E-06
rs207554348	ENSBTAG00000006505	-2.167642221	6.58E-13	1.41E-06
BovineNovelSNP579	ENSBTAG00000006505	1.45938953	2.68E-07	0.03867101
rs207554348	ENSBTAG00000006523	-1.931499631	1.62E-12	2.08E-06
rs207554348	ENSBTAG00000006904	-1.411825473	3.15E-07	0.04383373
BovineNovelSNP579	ENSBTAG00000006904	1.386856891	1.56E-08	0.00365732
rs207554348	ENSBTAG00000006921	-1.491312019	1.65E-09	0.00053688
rs207554348	ENSBTAG00000006990	-1.594446217	4.72E-09	0.00135396
BovineNovelSNP579	ENSBTAG00000006990	1.415834188	7.41E-09	0.00187453
rs207554348	ENSBTAG00000006999	-2.398780857	1.63E-13	1.04E-06
rs207554348	ENSBTAG00000007169	-1.404628699	8.84E-08	0.01617885
rs207554348	ENSBTAG00000007239	-2.37092338	1.93E-11	1.20E-05
BovineNovelSNP579	ENSBTAG00000007239	1.729466552	1.28E-07	0.02197581
rs207554348	ENSBTAG00000007901	-2.385945922	4.87E-12	4.68E-06
BovineNovelSNP579	ENSBTAG00000007901	1.713357773	9.71E-08	0.01745444
BovineNovelSNP2890	ENSBTAG00000007901	1.616494564	2.40E-07	0.03570202
rs207554348	ENSBTAG00000008428	-1.890181763	1.60E-11	1.05E-05
BovineNovelSNP579	ENSBTAG00000008428	1.44788656	2.13E-08	0.00481333

rs207554348	ENSBTAG00000008951	-2.939356878	8.00E-11	4.27E-05
BovineNovelSNP579	ENSBTAG00000008951	2.185292051	1.64E-07	0.02630587
rs207554348	ENSBTAG00000009773	-2.623061501	1.41E-12	1.97E-06
BovineNovelSNP579	ENSBTAG00000009773	1.904523571	3.32E-08	0.00701296
rs207554348	ENSBTAG00000010007	-1.647999901	4.96E-10	0.00019851
BovineNovelSNP579	ENSBTAG00000010007	1.261590211	2.13E-07	0.03219686
rs207554348	ENSBTAG00000010273	-2.951351717	6.53E-10	0.00025105
BovineNovelSNP579	ENSBTAG00000010273	2.285876242	1.78E-07	0.02817999
rs207554348	ENSBTAG00000011465	-1.537338538	2.30E-09	0.00071209
rs207554348	ENSBTAG00000011677	-1.506156566	1.63E-08	0.00378267
rs41890051	ENSBTAG00000012403	1.238042741	3.54E-07	0.04757233
rs207554348	ENSBTAG00000012638	-1.939834714	1.27E-13	1.04E-06
rs207554348	ENSBTAG00000012640	-2.212793251	3.49E-13	1.34E-06
rs207554348	ENSBTAG00000013201	-1.800388968	4.12E-12	4.17E-06
BovineNovelSNP579	ENSBTAG00000013201	1.320751766	4.29E-08	0.00864264
rs207554348	ENSBTAG00000013205	-1.798314718	6.45E-13	1.41E-06
rs207554348	ENSBTAG00000013290	-1.515991069	2.12E-09	0.00066801
BovineNovelSNP579	ENSBTAG00000013290	1.27669676	2.63E-08	0.00573749
rs207554348	ENSBTAG00000013356	-2.30086723	1.47E-07	0.02430409
rs472302580	ENSBTAG00000013356	3.076795609	6.51E-09	0.00176295
rs210747155	ENSBTAG00000013356	2.12095349	2.58E-07	0.03750555
rs207554348	ENSBTAG00000013368	-1.667697895	8.72E-13	1.68E-06
rs207554348	ENSBTAG00000013555	-1.583441339	1.96E-12	2.22E-06
rs207554348	ENSBTAG00000013706	-2.044338269	2.74E-11	1.55E-05
BovineNovelSNP579	ENSBTAG00000013706	1.498322459	1.35E-07	0.02258775
rs384819821	ENSBTAG00000014046	1.568367431	1.76E-08	0.00402639
rs207554348	ENSBTAG00000014122	-1.521315883	1.45E-11	1.01E-05
BovineNovelSNP579	ENSBTAG00000014122	1.199301124	6.69E-09	0.00176295
rs207554348	ENSBTAG00000014149	-2.183774501	1.44E-12	1.97E-06
rs207554348	ENSBTAG00000015592	-2.130792179	4.60E-13	1.34E-06
BovineNovelSNP2890	ENSBTAG00000015592	1.392672785	2.25E-07	0.03386956
rs207554348	ENSBTAG00000016566	-1.975258766	7.12E-12	6.45E-06
rs207554348	ENSBTAG00000017251	-1.747961965	1.97E-09	0.00063071
rs207554348	ENSBTAG00000017969	-1.724938488	2.26E-08	0.00504244
BovineNovelSNP579	ENSBTAG00000017969	1.490785807	8.58E-08	0.01585265
BovineHD0100032240	ENSBTAG00000017969	-1.013690614	3.34E-07	0.04553411
Hapmap55381-rs29025399	ENSBTAG00000017969	-1.013690614	3.34E-07	0.04553411
BovineNovelSNP2890	ENSBTAG00000018016	1.462232738	6.94E-09	0.00180319
BovineNovelSNP2931	ENSBTAG00000018016	1.695947861	3.20E-07	0.04431169
rs207554348	ENSBTAG00000018134	-1.87125535	4.89E-13	1.34E-06
BovineHD2300010006	ENSBTAG00000018134	0.926889293	1.05E-07	0.01832303
rs207554348	ENSBTAG00000018223	-1.634951796	1.47E-11	1.01E-05
rs207554348	ENSBTAG00000018280	-2.205365546	1.71E-10	8.23E-05
BovineHD0100032240	ENSBTAG00000018280	-1.165887524	2.51E-07	0.03682879
Hapmap55381-rs29025399	ENSBTAG00000018280	-1.165887524	2.51E-07	0.03682879
rs207554348	ENSBTAG00000018446	-1.778135936	2.96E-10	0.0001353
BovineNovelSNP579	ENSBTAG00000018446	1.413283002	4.32E-08	0.00864264
rs207554348	ENSBTAG00000019330	-2.225861293	1.11E-11	8.51E-06
BovineNovelSNP579	ENSBTAG00000019330	1.731805352	9.26E-09	0.00228175

rs207554348	ENSBTAG00000019627	-2.754891323	7.23E-09	0.00185375
rs207554348	ENSBTAG00000019669	-1.962757392	1.18E-11	8.72E-06
rs207554348	ENSBTAG00000020257	-1.741603216	3.56E-10	0.00015204
BovineNovelSNP579	ENSBTAG00000020257	1.473138854	4.85E-09	0.00137167
rs207554348	ENSBTAG00000020430	-1.31476908	3.67E-08	0.00759168
BovineNovelSNP579	ENSBTAG00000020430	1.140206458	1.20E-07	0.02072231
rs207554348	ENSBTAG00000020580	-2.101469515	2.60E-11	1.52E-05
BovineNovelSNP579	ENSBTAG00000020580	1.534419398	1.48E-07	0.02430409
rs207554348	ENSBTAG00000020676	-2.390294251	1.34E-12	1.97E-06
BovineNovelSNP579	ENSBTAG00000020676	1.663538755	1.35E-07	0.02258775
rs207554348	ENSBTAG00000021240	-2.034257171	5.48E-09	0.0015059
rs207554348	ENSBTAG00000021887	-1.82975457	2.96E-08	0.00632867
rs207554348	ENSBTAG00000022779	-1.592503178	3.55E-10	0.00015204
rs207554348	ENSBTAG00000023648	-1.41688526	1.06E-08	0.00257986
BovineNovelSNP2890	ENSBTAG00000023648	1.142713724	1.60E-07	0.02579907
BovineNovelSNP579	ENSBTAG00000023648	1.172708212	1.80E-07	0.02817999
rs207554348	ENSBTAG00000031647	-3.923999467	3.29E-10	0.00014712
rs207554348	ENSBTAG00000031950	-2.16809336	1.90E-12	2.22E-06
BovineNovelSNP579	ENSBTAG00000031950	1.541753124	8.01E-08	0.01540411
rs207554348	ENSBTAG00000037826	-1.382038921	6.62E-09	0.00176295
BovineNovelSNP579	ENSBTAG00000037826	1.156821332	8.48E-08	0.01583543
rs207554348	ENSBTAG00000039556	-1.566334549	2.50E-12	2.67E-06
BovineNovelSNP579	ENSBTAG00000039556	1.125960225	6.47E-08	0.01269327
rs207554348	ENSBTAG00000046152	-1.93380839	7.86E-10	0.0002763
BovineNovelSNP579	ENSBTAG00000046152	1.530985893	1.01E-07	0.01790787
rs207554348	ENSBTAG00000046158	-2.650342792	2.88E-10	0.00013523
BovineNovelSNP2890	ENSBTAG00000046158	2.013602415	7.17E-08	0.01391896
BovineNovelSNP9232	ENSBTAG00000046158	1.922574924	1.79E-07	0.02817999
rs207554348	ENSBTAG00000048720	-2.994552213	8.47E-08	0.01583543
rs207554348	ENSBTAG00000048737	-1.894842052	7.90E-10	0.0002763
rs1118122708	ENSBTAG00000048737	-1.312587991	3.00E-07	0.04293916
rs207554348	ENSBTAG00000048835	-3.655440782	1.48E-09	0.00049203
rs207554348	ENSBTAG00000049416	-1.912527711	1.64E-11	1.05E-05
BovineNovelSNP579	ENSBTAG00000049416	1.343541695	3.74E-07	0.04998529
rs207554348	ENSBTAG00000049808	-2.407355462	7.26E-10	0.00026702
BovineNovelSNP579	ENSBTAG00000049808	1.901195599	1.03E-07	0.018162
rs211132440	ENSBTAG00000050072	-1.838155096	1.22E-08	0.00292783
rs207554348	ENSBTAG00000050072	-1.464703791	3.02E-07	0.04293916
rs210238564	ENSBTAG00000050072	-1.84720955	1.35E-07	0.02258775
rs207554348	ENSBTAG00000050618	-2.767330452	8.62E-09	0.00215338
rs207554348	ENSBTAG00000051132	-3.573898428	2.83E-09	0.00082507
BovineNovelSNP579	ENSBTAG00000052465	1.190798997	8.14E-08	0.01549265
rs207554348	ENSBTAG00000052465	-1.507322777	5.32E-10	0.00020891
rs207554348	ENSBTAG00000053557	-2.324795623	7.38E-12	6.45E-06
BovineNovelSNP10422	ENSBTAG00000053557	1.844810697	3.06E-07	0.04321986
rs207554348	ENSBTAG00000054765	-1.428232382	2.33E-08	0.0051392
BovineNovelSNP579	ENSBTAG00000054765	1.180992818	3.53E-07	0.04757233
rs207554348	ENSBTAG00000054882	-3.450762956	5.22E-09	0.00145431

Table 5.S10 Common differential expressed genes between the current study and those identified by Tizioto et al. for animals who were challenged by different pathogens

Challenge pathogen	Common DE genes between two studies
BRSV	*LRG1, RAB20, CFB, CD163, SOD2, ALPL, S100A9, NOL3, DYSF, TCN1, S100A8, CHI3L1, UPP1, S100A12, HP, WIPI1, HIST1H1C, TGM3, MMP9, IL12B, MYBPH, DEFB4A, NUPR1, RETN, PGLYRP1, BPI*
IBR	*LRG1, RAB20, CFB, CD163, SOD2, ALPL, S100A9, NOL3, DYSF, TCN1, S100A8, CHI3L1, UPP1, GPR84, IL1R2, S100A12, TNFAIP6, PTPN5, HP, WIPI1, HIST1H1C, TGM3, MMP9, RBPMS2, IL12B, ANKRD22, P2RX1, MYBPH, HTRA1, DEFB10, DEFB4A, NUPR1, RETN, PGLYRP1, BPI*
BVDV	*BMX, LRG1, RAB20, CFB, CD163, S100A9, NOL3, DYSF, S100A8, CHI3L1, GPR84, THY1, S100A12, ABCA6, MAPK13, HP, WIPI1, BST1, TGM3, MMP9, RBPMS2, GPC3, HTRA1, ITGA9, DEFB10, NUPR1, RETN, PGLYRP1, REEP1*
M. haemolytica	*BMX, LRG1, CFB, SHROOM4, KREMEN1, CD163, SOD2, ALPL, S100A9, ACVR1B, NOL3, TCN1, S100A8, CHI3L1, UPP1, IL1R2, S100A12, RAB3D, ABCA6, MAPK13, HP, WIPI1, IL1RAP, ALOX5AP, HIST1H1C, OSCAR, MGAM, BST1, TGM3, MMP9, MYBPH, HTRA1, ITGA9, DEFB10, OLFM4 RETN LTF PGLYRP1 REEP1*
P. multocida	*BMX, CA4, CD163, S100A9, S100A8, CHI3L1, S100A12, TNFAIP6, HP, WIPI1, TGM3, MMP9, RBPMS2, GPC3, HTRA1, DEFB10, OLFM4, RETN, LTF, PGLYRP1*
M. bovis	*LRG1, S100A9, S100A8, MGAM, TGM3, MMP9, HTRA1, PGLYRP1*
All pathogens	*S100A8, S100A9, MMP9, TGM3, PGLYRP1*

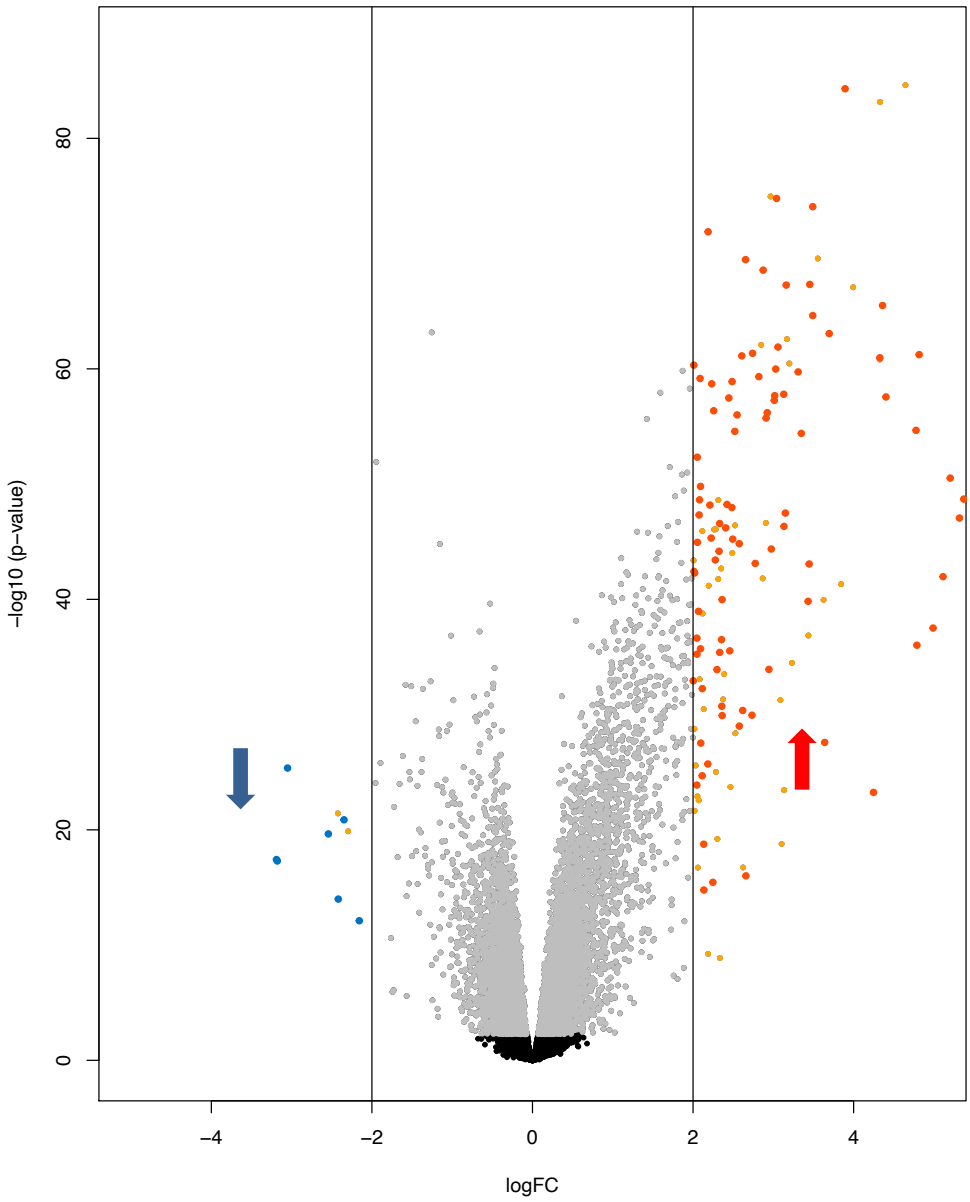

Figure 5.1 Volcano plot of 7 down-regulated genes (blue) and 94 up-regulated genes (red) for the BRD animals

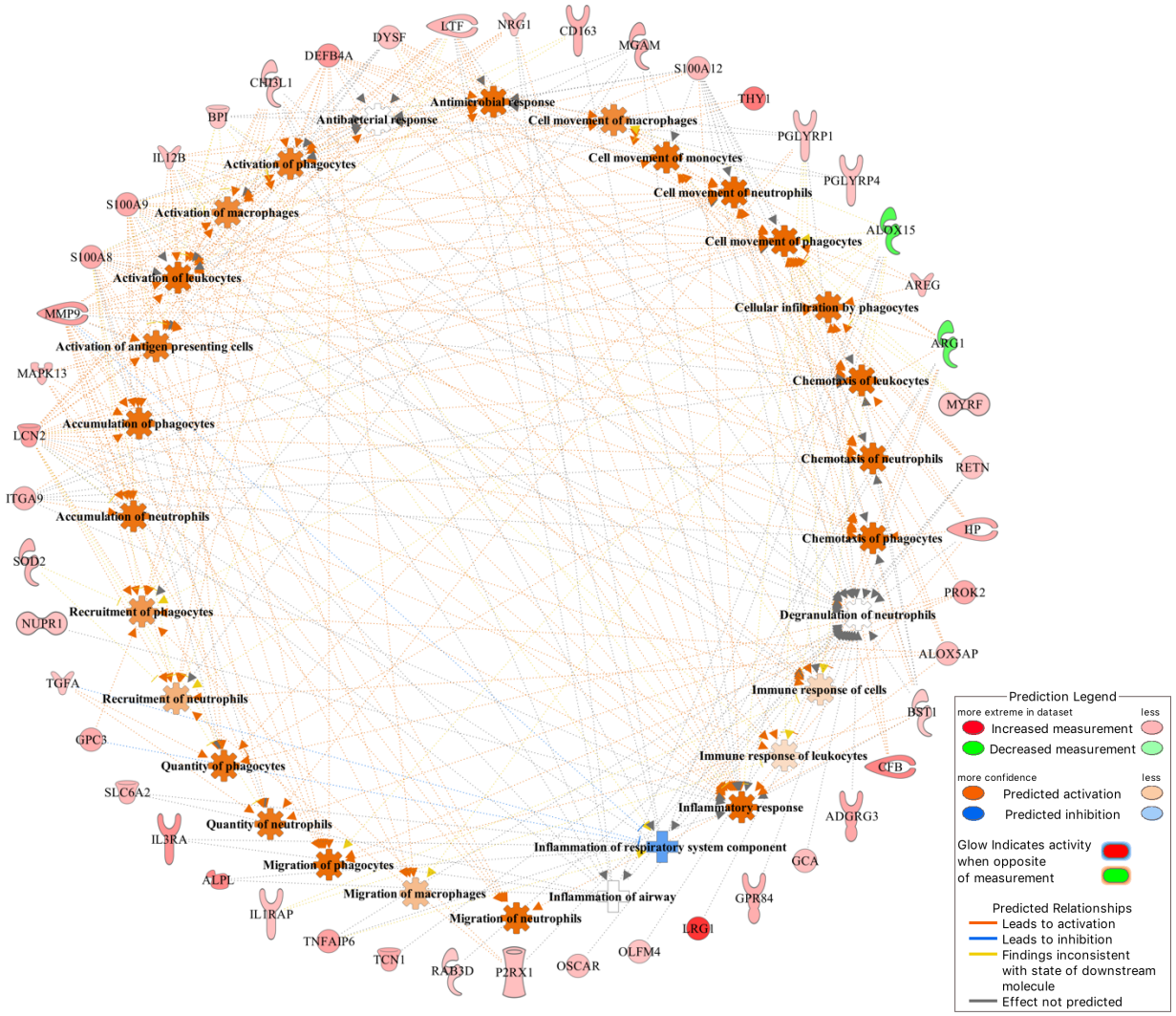

Figure 5.2 Gene regulatory network of inflammatory response

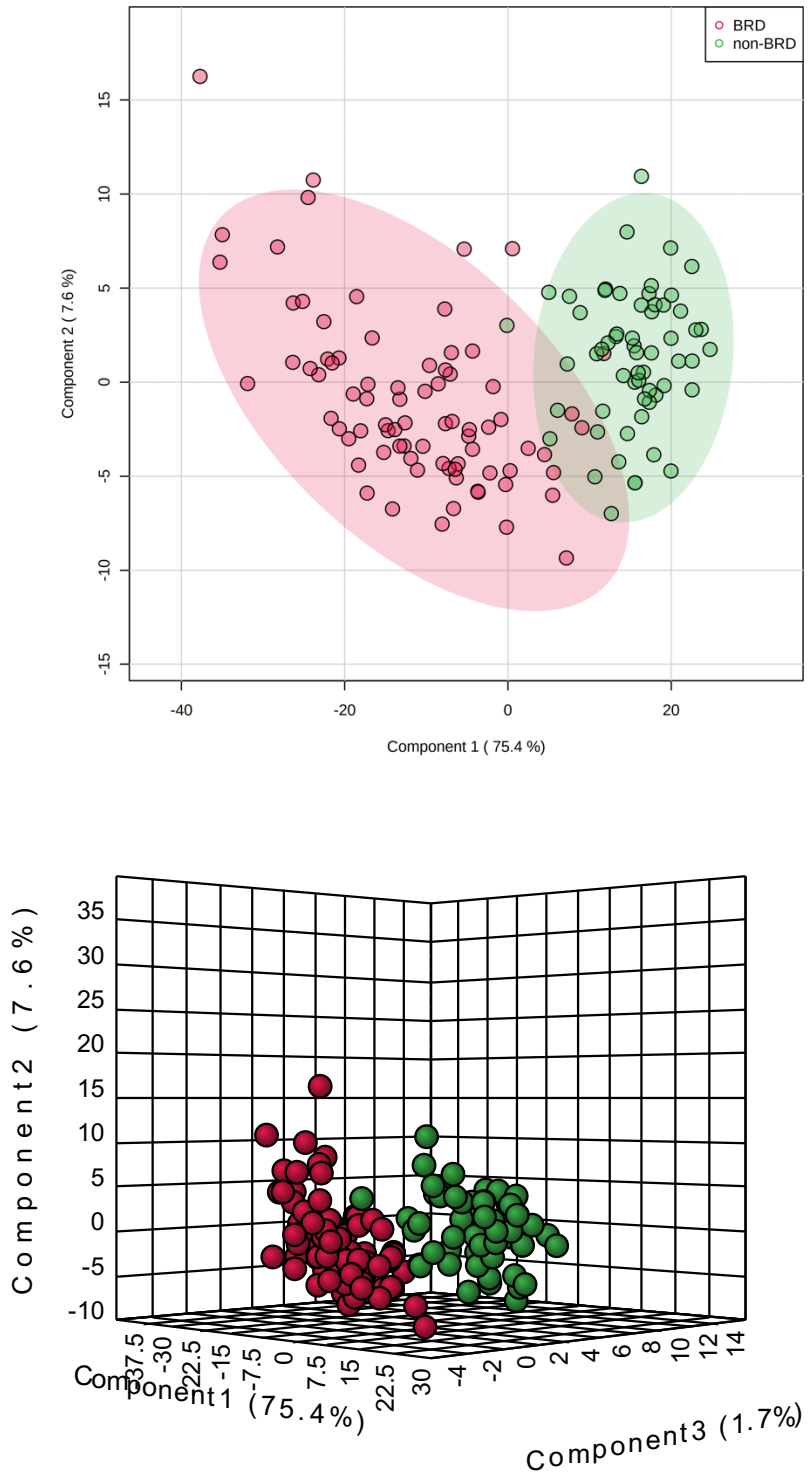

Figure 5.3 The 2D and 3D PLS-DA plot of animals in BRD (red) and non-BRD (green) group based on the expression of 101 DE genes

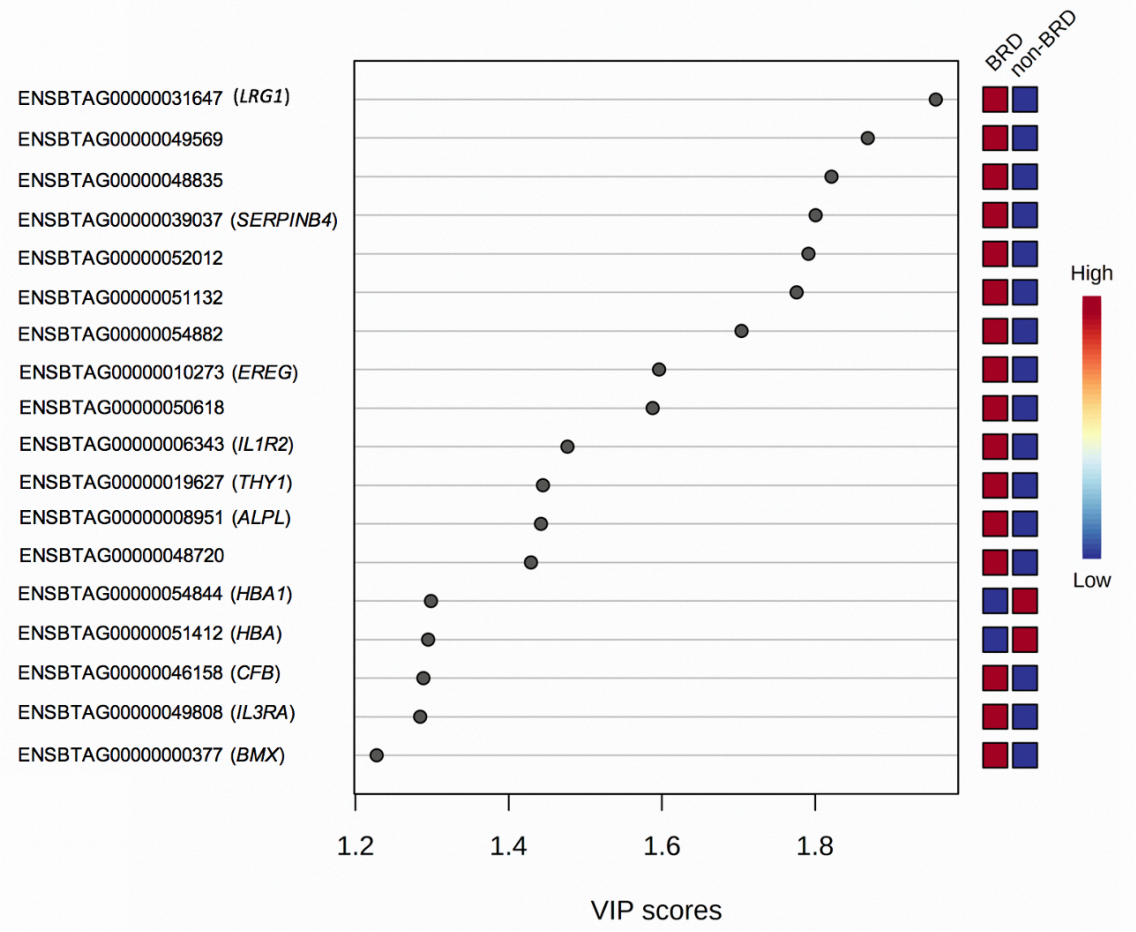

Figure 5.4 The DE genes with VIP score > 1.2 in PLS-DA

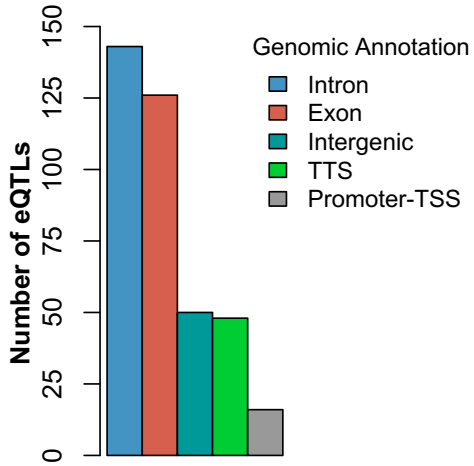

Figure 5.5 Distribution of eQTLs on chromosomes

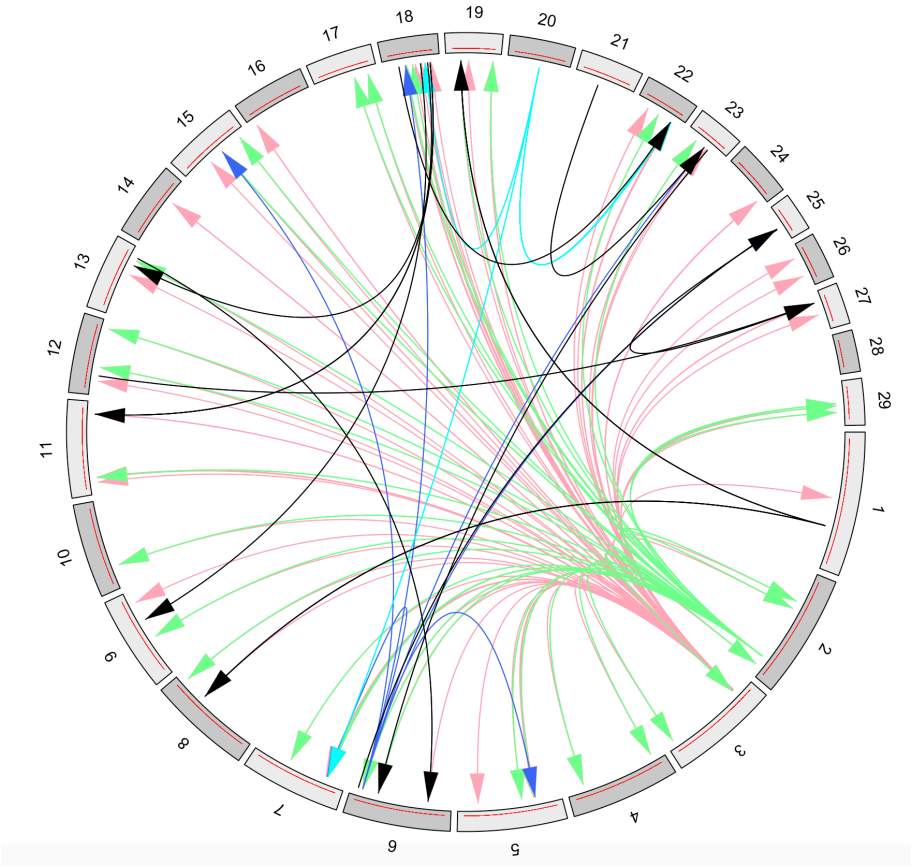

Figure 5.6 Circos plot showing trans eQTLs and their relationships to differentially expressed genes

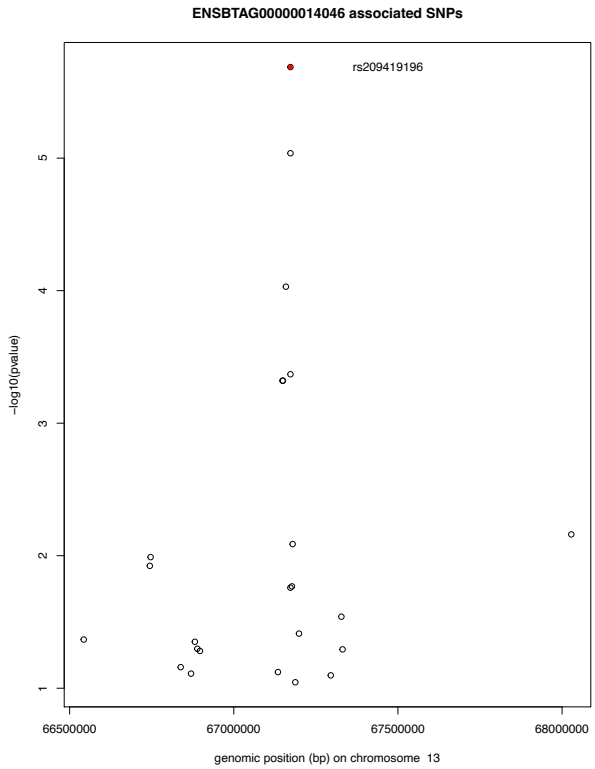

Figure 5.7 Regional Manhattan plot for all SNPs around 1Mbp up- and down-stream of *BPI*

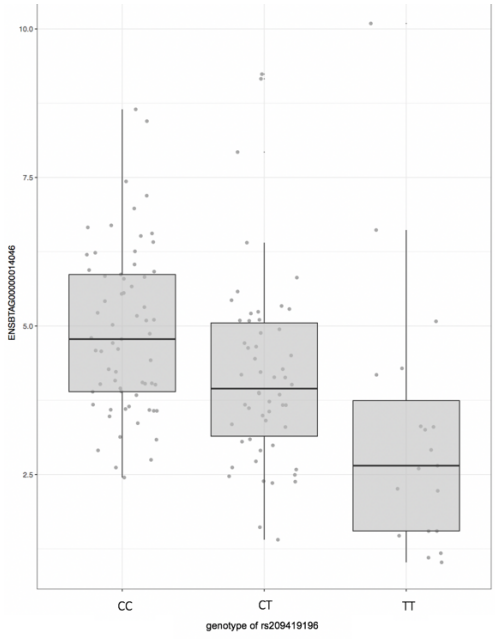

Figure 5.8 The box plot of effects of three genotypes (CC, CT, and TT) of rs209419196 on the expression of *BPI*

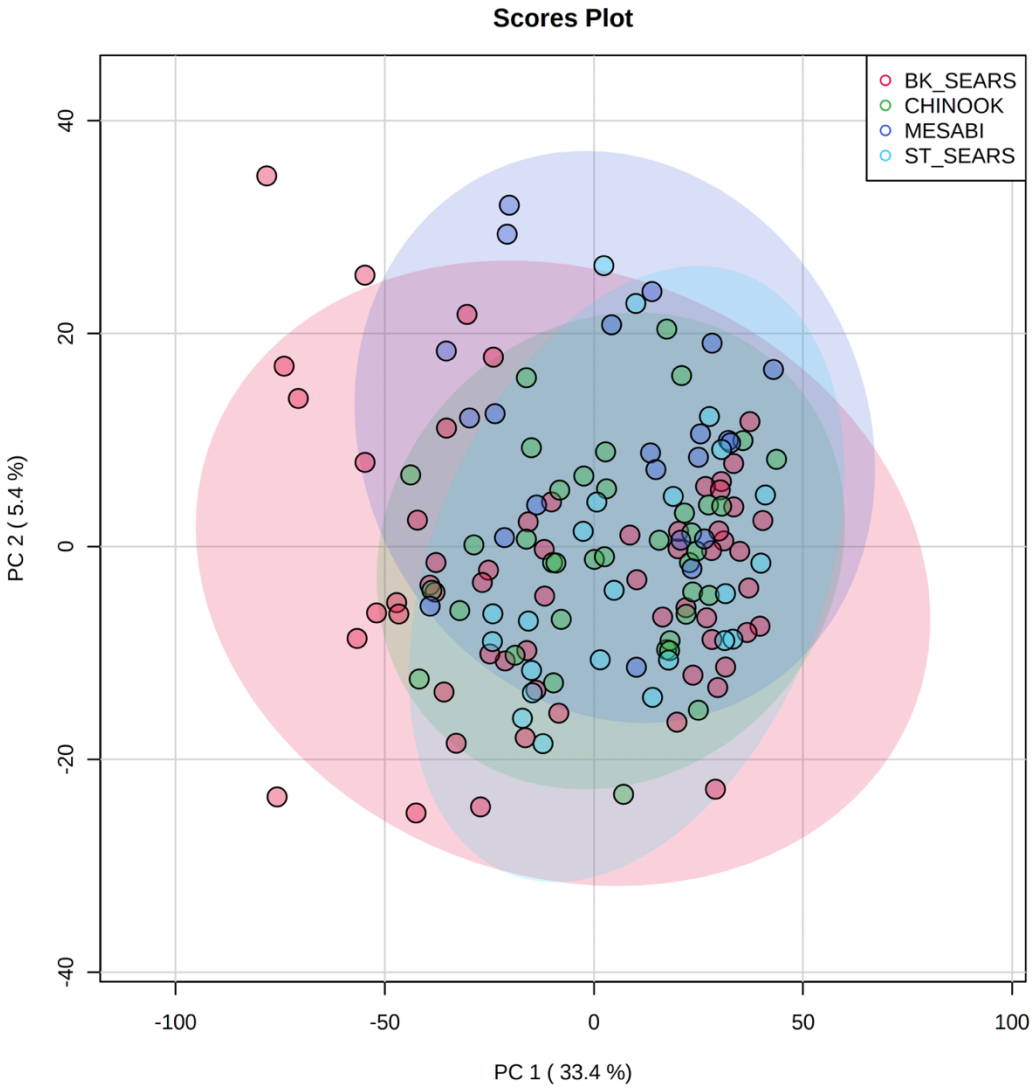

Figure 5.S1 The principal component analysis plot of animals at four feedlots based on all expressed genes

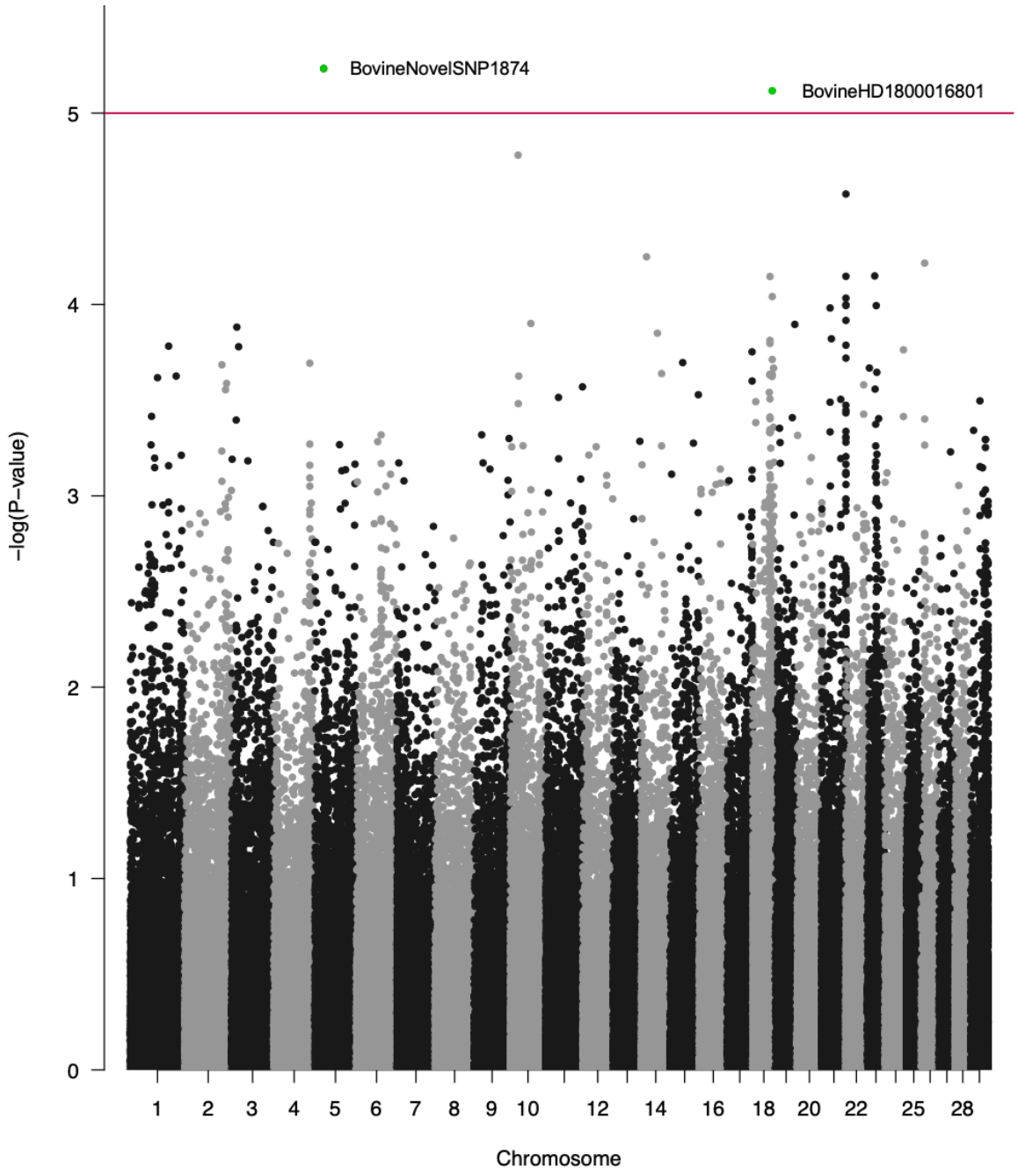

Figure 5.S2 Manhattan Plot of GWAS

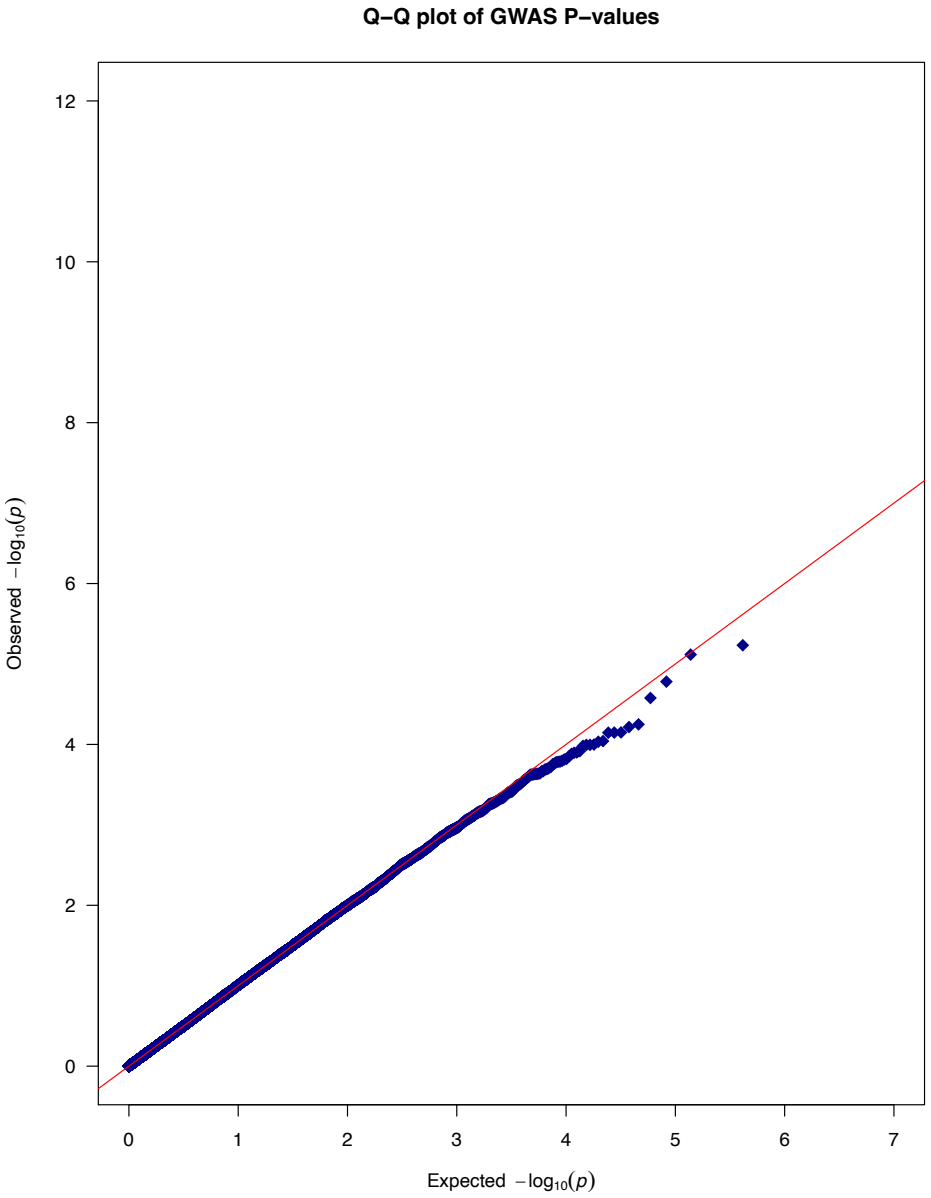

Figure 5.S3 Q-Q Plot of *P*-values

Chapter 6. General discussion

Feed efficiency, carcass merit and resistance to bovine respiratory disease (BRD) are key economically relevant traits related to the profitability and sustainability of beef production. Breeding and selecting animals with desirable traits is one of the goals of beef cattle genetic improvement programs. In the past decade, genome-wide association studies (GWAS) have been widely used to study the association between genetic variants and the variation in beef cattle traits. However, for most SNPs or genes identified as associated with these important traits, our understanding of which genetic variants drive the phenotypic variation and how genetic variants lead to phenotypic variation is still limited. In addition, knowledge is accumulating regarding the association between different omics (e.g., transcriptomics and metabolomics) and the variation of these traits as well as the interrelationship between different omics layers involved in the phenotypic variation. Improving the understanding of these questions could provide a holistic view of the genetic and molecular background of the studied traits. Therefore, in this thesis, integrative analyses of multi-omics data were performed to give insights into these questions. This chapter consists of key findings of each study and their implications. Additionally, the limitations of current studies and recommendations for future research are discussed.

6.1 Summary of key findings and implications

Study I. Early studies in human have suggested that metabolite concentrations (direct readouts of biological processes) could provide functional links between genetic variance and external phenotypes in GWAS (Suhre and Gieger, 2012). However, in beef cattle, the genetic architecture of blood metabolites is largely unknown. Therefore, the first step in this work was to study the genetic architecture of blood metabolites in beef cattle (Chapter 2), which also provided the prerequisite knowledge for subsequent studies that link blood metabolites with genetic variants,

feed efficiency, and carcass merit traits (Chapter 3 and Chapter 4). The first study (Chapter 2) estimated heritability for blood metabolites and identified genomic regions that were associated with variation of metabolites using 50K genotypes. In this study, we found multiple genomic regions associated with metabolites and observed some genomic regions with pleiotropic effects on two or more metabolites. Although the polygenic and pleiotropic nature of the metabolite variation have been previously reported in human metabolomics studies (Hu et al., 2018; Gallois et al., 2019), this study is the first attempt to characterise the genetic basis of blood metabolites in crossbred beef cattle. The results obtained in this study could be used as baseline information for further genetic research using blood metabolites. Blood metabolites have previously been reported to be associated with feed efficiency and carcass merit traits (Karisa et al., 2014; Connolly et al., 2019). Therefore, there is an opportunity to identify functional SNPs or genes and molecular mechanisms associated with feed efficiency and carcass merit traits by incorporating metabolites into genetic studies (Chapter 3 and Chapter 4).

Study II&III. In Chapter 3 and Chapter 4, multiple metabolites were found to be associated with feed efficiency and carcass merit traits, which linked the variation at the phenotypic level with variation at the metabolomic level. According to the comparison of our results with those from previous studies (Karisa et al., 2014; Jorge-Smeding et al., 2019; Foroutan et al., 2020), we found that the results were in good agreement. The consistency of results from different studies suggests that metabolites have the potential to be used as biomarkers for feed efficiency and carcass merit traits. Additionally, by combining the results of metabolome-genome wide association studies for trait-associated metabolites using whole genome sequence variants, functional SNPs and candidate genes associated with these traits were identified. By comparing our results with previous GWAS for feed efficiency and carcass merit traits (Seabury et al., 2017;

Wang et al., 2020; Zhang et al., 2020a), we performed an integrative analysis of genomic and metabolomic data to give more insights into the above-mentioned questions: 1) which genetic variants drive the phenotypic variation; 2) how genetic variants lead to phenotypic variation, and 3) the interrelationship between different omics layers involved in the phenotypic variation. For example, Zhang et al. (2020a) identified *ADGRF1*, *ADGRF5*, *GTPBP8* and *NEPRO* as associated with both RFI and dry matter intake (DMI), and the same genes were identified to affect these same traits in this study. Further, the results of this study indicated that L-tyrosine was associated with *ADGRFI* and *ADGRF5* which might explain the associations of *ADGRF1*, *ADGRF5* with RFI and DMI. As for *GTPBP8* and *NEPRO*, both genes were associated with lysine, a metabolite associated with both RFI and DMI in this study. Therefore, we speculated that the overlap of genes found in previous GWAS indicates that these genes may be functional genes that cause phenotypic variation by regulating the synthesis and degradation of associated metabolites. Furthermore, Chapter 3 and Chapter 4 provided additional insights into the potential biological processes of traits of interest. Several key metabolic processes including metabolism of lipids, molecular transportation, and carbohydrate metabolism were identified, and their related gene networks were constructed in our studies. Identification of these biological processes and their corresponding genes could help prioritize candidate genes for identification of functional or causal gene polymorphisms responsible for the phenotypic variation, and to identify and optimize molecular biomarkers (Sun and Hu, 2016; Lee et al., 2019). Therefore, the studies in Chapter 3 and Chapter 4 have shown that metabolites play important roles in the variation of both feed efficiency and carcass merit traits. Integration of metabolomic and genomic data could help to identify functional or causal SNPs or genes and interpret the biological meaning of the candidate genes identified in GWAS. Thus, our findings have broadened the knowledge on the genetic and molecular

mechanisms underlying variation in beef cattle traits and supports future application of multi-omics approach to beef cattle research.

Study IV. Previous research studies have indicated that host genetics plays a significant role in regulating immune response and determining susceptibility and resistance to bovine respiratory disease (Muggli-Cockett et al., 1992; Snowder et al., 2006; Snowder, 2009; Emam et al., 2019). For example, heritability estimates (0.07 - 0.29) of BRD susceptibility have been reported in different cattle populations (Snowder et al., 2005; Schneider et al., 2010; Neibergs et al., 2014a, 2014b). In Chapter 5, we observed that the heritability of BRD susceptibility in our studied population was 0.43 ± 0.51. Even though this estimate had a relatively large standard error that may have resulted from the limited number of animals utilized in this study, this result reveals potential for genetic variability of BRD susceptibility in beef cattle which could be exploited to breed for BRD resistant or resilient animals. In this study, two significant SNPs (BovineNovelSNP1874 on chromosome 5 and BovineHD1800016801 on chromosome 18) were identified as associated with BRD susceptibility by GWAS. Interestingly, the most significant SNP (BovineNovelSNP1874) explained 17% of the phenotypic variance for BRD susceptibility. This implies that this SNP could be a major quantitative trait nucleotide (QTN) or in linkage disequilibrium with a major quantitative trait locus (QTL) for BRD susceptibility in the studied population. However, the proportion of phenotypic variance explained by significant SNPs could be overestimated because of the limited number of animals used. Thus, future research utilising larger sample size is warranted to more accurately estimate the heritability of BRD susceptibility. In addition to the genomic level, many transcriptomic studies have reported many differentially expressed genes (DE genes) between BRD and non-BRD animals (Tizioto et al., 2015; Scott et al., 2020; Sun et al., 2020; Jiminez et al., 2021), indicating that gene expression is an important

intermediate phenotype associated with BRD. In this study, we profiled the transcriptomic signature of animals with and without BRD and identified 101 genes differentially expressed between BRD and non-BRD animals from four feedlots. Most of the DE genes identified in our study have been reported several times in previous transcriptomics studies for BRD. For example, *CFB* is an important DE gene identified in this study, which is also a DE gene reported in all BRD transcriptomic or single pathogen challenge studies (Nalpas et al., 2013; Tizioto et al., 2015; Johnston et al., 2019; Scott et al., 2020, 2021; Sun et al., 2020; Jiminez et al., 2021). This indicates that certain genes have the potential to be used as biomarkers for BRD diagnosis. By analyzing transcriptomic data, we identified 18 DE genes that had good predictive ability for BRD status. We further applied these 18 DE genes to populations from any of four feedlots and they all had high accuracy and reliability to predict BRD status. Therefore, we suggest that applying these 18 DE genes to an independent population or practical feedlot beef production to test their feasibility and accuracy. Finally, we investigated the association between genotype and gene expression through expression QTL (eQTL) analysis. Multiple cis-eQTLs and trans-eQTLs associated with the expression of DE genes were identified. Interestingly, the most significant SNP (BovineNovelSNP1874) associated with BRD susceptibility was identified as a cis-eQTL for the DE gene *GPR84*, indicating that this SNP may be causal for BRD susceptibility through its cis effects on the expression of *GPR84.* Additionally, Neibergs et al. (2014) reported a genomic region including the gene *BPI* as associated with BRD susceptibility in Holstein calves, and suggested that variants within or in close proximity to this gene have functional relevance in modulating susceptibility to BRD in cattle. *BPI* was differentially expressed between BRD and non-BRD animals in this study and as well as those of previous studies (Tizioto et al., 2015; Johnston et al., 2019, 2021; Jiminez et al., 2021). Through combing transcriptomic data and genomic data in this

study, we further identify the most possible causal SNP among all variants within or in close proximity to the gene *BPI.* Although these causal SNPs still need further investigation and validation in an independent population, our analysis gives additional information of causal SNP identification, GWAS interpretation and potential molecular mechanisms. Therefore, this study provides further enrichment of our understanding of the genetic regulation of gene expression associated with BRD in beef cattle. This information could not be captured by using GWAS or transcriptomic analysis alone, which exhibits the advantage of multi-omics study over single omics study.

6.2 Limitations and recommendations

6.2.1 Application of functional SNPs in genomic prediction

Genomic selection has revolutionized animal breeding and many traits have benefitted from this technology. The accuracy or reliability of the prediction of the genomic merit of the candidates is one of the major factors that affect the progress rate of genetic improvement (Georges et al., 2018). However, the prediction accuracy for beef cattle traits such as feed efficiency, carcass merit, and resistance to BRD remains relatively low (Mujibi et al., 2011; Bolormaa et al., 2013; Akanno et al., 2014; Lu et al., 2016; Silva et al., 2016). The lower prediction accuracy of beef cattle may be due to the lack of a large number of reference populations within a breed and insufficient understanding of genetic mechanisms of complex traits. The low accuracy makes it difficult to identify beef cattle with high breeding values, which affects the rate of genetic gain.

As reviewed in Chapter 1, including biological or functional information or causal DNA variants into the SNP panels or in the statistical models could improve the accuracy of predicting genetic merit of selection candidates (Snelling et al., 2013; MacLeod et al., 2016; Fang et al., 2017b). This method could improve prediction accuracy, especially in across breed prediction, or

reduce the marker density used in genomic prediction while retaining accuracy (Melzer et al., 2013; MacLeod et al., 2016; Sarup et al., 2016; Fang et al., 2017a; Gebreyesus et al., 2019), which could improve the rate of genetic gain in beef cattle and reduce the cost of genotyping services. Additionally, the knowledge about genomic influence and molecular background of these traits may also help design more effective genomic selection strategies to improve the prediction accuracy, and the rate of response to genomic selection for the traits of interest (Brandebourg et al., 2013).

Findings in this thesis have broadened our knowledge on the genetic background of feed efficiency, carcass merit and BRD resistance. Also, this thesis identified and refined many functional SNPs and genes associated with these traits by analyzing multi-omics data. Due to time and funding issues, we are not able to develop a SNP evaluation panel that includes these functional SNPs or SNPs harboured in functional genes. It is expected that such functional SNP panel could improve the prediction accuracy of genomic prediction in commercial crossbred beef cattle. In addition to developing SNP evaluation panels, these functional SNPs can be weighted in statistical models or used in more advanced predictive models such as genomic feature BLUP (Fang et al., 2017a) or Bayes RC (MacLeod et al., 2016). Overall, we recommend implementing further projects aimed at improving the accuracy of genome predictions by including the variants identified in order to elucidate this hypothesis.

6.2.2 Utilization of comprehensive metabolomic and transcriptomic profile

This study profiled the plasma metabolome from frozen blood samples of crossbred or multi-breed beef cattle using nuclear magnetic resonance methodology and 31 metabolites were quantified and used in the analyses. The results showed that individual metabolites only account for a small proportion of phenotypic variance. This relatively small percentage of phenotypic

variance likely indicates that the complex traits are regulated by multiple metabolic pathways involving many metabolites with each having only a small contribution. However, it is also possible that we were not able to identify metabolites with major influences on the studied traits due to the limited number of metabolites that was profiled. Carmelo et al. (2020) examined the association between 729 blood metabolites and RFI in pigs, many more metabolites (102 metabolites for Duroc and 73 metabolites for Landrace, respectively) were identified as associated with RFI. The proportion of associated metabolites ranged from 10% to 14% of all profiled metabolites, which is similar to the proportion of 16% from this study. However, the phenotypic variance explained by metabolites was not reported in the study of Carmelo et al. (2020). In addition, other metabolomics studies in cattle also reported some metabolites are associated with feed efficiency which are not quantified in the current study (Clemmons et al., 2017; Novais et al., 2019; Wang and Kadarmideen, 2019). Therefore, we anticipate that there will be more metabolites related to feed efficiency and carcass traits in beef cattle, and it is recommended to conduct a comprehensive study that includes a more complete metabolite profile, e.g., untargeted analysis (Zhao and Li, 2020). We analyzed plasma metabolites because blood is the major highway for absorption and transportation of nutrients to the different organs and tissues, and metabolites carried by blood are directly involved in metabolic processes as substrates or products. Also, blood is easy to collect from animals, which makes blood biomarkers more practical in application. We collected blood samples at one time point that is the first day of feeding test. This time point could represent the animals' metabolism in the feeding test period since the metabolite profile could show a level of temporal stability and predictability (Carmelo et al., 2020). However, it is not able to represent the situation in a lifetime. Karisa et al. (2014) reported the metabolites associated with feed efficiency varied over time, although some of them are common. We recommend that future

studies should consider collecting blood samples from animals at multiple developmental stages or from animals in different environments, which may help to identity more accurate biomarkers. This also prompts a hypothesis that different genes may play a role in metabolism at different stages of life. In addition to blood, more and more evidence shows that there are other important organs/tissues (such as liver, adipose tissue, skeletal muscle, and the digestive tract) in the body whose metabolism potentially influences feed efficiency and carcass merit traits of the animal (Alexandre et al., 2015; McKenna et al., 2018, 2021; Freetly et al., 2020). Therefore, to give additional insights into the relationship between animal metabolism and phenotypes, further metabolomics studies should be pursued considering these organs and tissues. This information may be useful for further gene prioritization.

In the transcriptomic analysis of BRD (Chapter 5), there are also limitations on sampling source and time. We only collected the blood samples from sick and healthy animals once within 50 days on feed after cattle arrived at feedlots. In addition, the batch effect of RNA sequencing caused by technical differences between samples, such as the type of sequencing machine or even the technician that ran the sample, also needs to be considered to because it could lead to an unfavourable impact on downstream biological analysis. The normalization methods, such as transforming the raw counts to (logarithms of) CPM, TPM or RPKM/FPKM, the trimmed mean of M values (TMM) (Robinson and Oshlack, 2010), or relative log expression (RLE) (Risso et al., 2014) could correct the differences in the overall expression distribution of each sample across batches (Zhang et al., 2020b). For differential expression, many common methods or procedures, such as edgeR (Robinson et al., 2009) and DESeq2 (Love et al., 2014), suggest to include batch variables as covariates in the linear models behind these methods to account for the impact of batch. In the study of this thesis, we used the TMM normalization method and fitted batch effects

into the linear model to avoid the potentially existing batch effect in the analysis. Also, the principal component analysis (PCA) and gene differential expression analysis were performed, and results do not show a significant batch effect. However, future studies may include more animals from different sources and samples could not be processed at once, this batch effect should always be considered.

6.2.3 Precision of molecular mechanisms and biomarkers

Incorporating metabolomic data or transcriptomic data into genetic studies could offer insights into molecular architecture and biological background of complex traits. However, the precise molecular mechanisms are not very clear. For example, we found that eQTL SNPs are mainly located in the intronic and exonic regions. These eQTL SNPs within genes may affect gene expression through multiple mechanisms including amino acid sequence changes that may inhibit or stop gene transcription, as well as exonic splicing enhancers/silencers and intron splice enhancers/silencers that could regulate alternative splicing (Wang et al., 2009; Cooper, 2010; Kreimer and Pe'er, 2013; Deng et al., 2017). However, eQTL SNPs within genes were also reported to play a role in long-range gene regulation (Jowett et al., 2010; Ragvin et al., 2010; Kreimer and Pe'er, 2013). Therefore, some eQTL may influence the expression of remote genes at distance, rather than the expression of those genes that host them, and this regulatory pattern was also observed for some eQTLs in this study. In addition to those eQTLs located in the intronic and exonic regions, some eQTLs were in promoter-transcription start site (TSS), and transcription termination site (TTS). For example, rs209419196 was identified as a cis-eQTL associated with the expression of *BPI*, and it was predicted to be in the promoter of the transcript of *BPI*. However, the precise molecular mechanisms underlying most eQTLs, especially those in non-coding regions (distal regulatory regions), are still elusive. To further annotate these eQTLs, more precise and

comprehensive genomic information (annotation) of these distal regulatory regions is needed in cattle, and this is one of the core aim of the Functional Annotation of Animal Genomes (FAANG) project (Andersson et al., 2015). Our genomic, transcriptomic and metabolomic data have been used in the international BovReg project (https://www.bovreg.eu/) whose outcomes will be integrated into FAANG. It is expected that the research results from the FAANG project will lead to a better SNP functional annotation and increase our understanding of genomic structure and function. Additionally, epigenetics studies including DNA methylation via methylation sequencing (Methyl-Seq) and chromatin modification studies through chromatin immunoprecipitation sequencing (ChIP-Seq) and transposase-accessible chromatin sequencing (ATAC-Seq) (Dirks et al., 2016) may further aid in identification of genomic transcription regulatory regions in beef cattle.

In this thesis, some molecular biomarkers were proposed, for example, lysine and L-tyrosine for RFI, creatinine for hot carcass weight and rib eye area, and 18 DE genes for BRD. The literature or biomarker analysis indicates that these molecular biomarkers may be useful to select superior animals or diagnose disease. In the BRD study, we also collected blood metabolites, but the data has not been analyzed due to time constraints. Although the predictive ability of transcriptomic biomarkers is excellent, it is expected that the performance of prediction of phenotype by fitting both metabolomic and transcriptomic biomarkers to a prediction model would be better than using any single one of them.

6.2.4 Multi-omics data integration in future research

As sequencing technologies continue to improve and collect more in-depth data, and as the cost of sequencing continues to fall, researchers are trying to integrate and piece multi-omics data in different formats together to study complex biological questions. Huang et al. (2017) suggests

that “more is better” in multi-omics studies. However, adding omics datasets for the sake of adding more data might not always be a good idea, because they may carry useless or unrelated information that leads to more noise, redundancy and an increased computational burden in the multi-omics studies (Picard et al., 2021). Inappropriate models and incorrect integration methods in multi-omics analysis could even lead to worse performance (Ma et al., 2016; Picard et al., 2021). Additionally, for a multi-omics study, it is apparent that the costs and labour increase as more omics data and sample size are included. Therefore, we recommend choosing the most suitable tissues and most important omics layers with limited research funding. For example, multiple organs (e.g., liver and rumen) and tissues (e.g., muscle and adipose tissues) are involved in the variation of RFI. In this study, we used the blood sample that is easily accessible and could represent whole body metabolism to study the metabolites associated with RFI. These blood biomarkers are also more practical in production than those in liver or rumen. However, if the research topic is liver metabolism, the samples from liver will be more appropriate. In addition to collect field data, currently, many publicly available data sources could be utilized to serve multi-omics analysis, but the data from different sources may introduce new biases in the data analysis, which should be kept in mind when looking for the most appropriate data sources.

Statistical integration and network-based integration are two main data integration approaches. In our study, we mainly applied the statistical methods to identify the association between molecules at different omics layers and then draw inference from the associations. However, the association between different omics levels could provide evidence of a possible relationship between them, but the causality between omics levels is unknown. For example, when a molecular trait is associated with a phenotypic trait, there may be at least three possible explanations for the association: the molecular trait influences the phenotypic trait; the phenotypic

trait influences the molecular trait; or both are affected by a confounding factor. To overcome this limitation, we further performed genetic mapping for both molecular traits and phenotypic traits, which provides evidence of a relationship between genotype, molecular traits, and phenotypic traits through co-mapping. i.e., if a molecular trait and a phenotypic trait map to the same genomic region, then the molecular trait may be the link between genotype and phenotype. This method is straightforward, but it is difficult to identify the complex interaction between molecules at different omics layers. The network-based integration can integrate and use pre-existing biological knowledge in the database and combine it with experimental data to show the potential biological connection between molecules and facilitate the interpretation from a biological point of view and not from the data point of view (Yan et al., 2018; Zhou et al., 2020). This method could show the connection between different biomolecules and biological functions. It also could complement statistical data integration and help to integrate existing knowledge and interpret results (Zhou et al., 2020). Although network-based integrations are promising, it is worth noting that the prior information is often biased, which may consequently result in unreliable or unrelated connections in the network. Moreover, for the integration of more complex multi-omics datasets, some common challenges may arise, including dimension reduction, data heterogeneity, missing data, class imbalance and scalability issues (Mirza et al., 2019; Song et al., 2020; Picard et al., 2021). Deep learning architectures (Min et al., 2017; Kim et al., 2018) can better recognize complex features through representation learning with multiple layers and facilitate the integrative analysis through effectively addressing the common challenges (Mirza et al., 2019). The attempts of deep learning methods (e.g., based on artificial neural networks) have been applied in human studies (Sharifi-Noghabi et al., 2019; Lin et al., 2020). With the development of integration tools and the continuous growth of data integration experience, we recommend using these more advanced

integration methods to study multi-omics data for complex traits in future beef cattle genetic research.

6.3 Overall summary

Many SNPs and candidate genes have been identified to be associated with feed efficiency, carcass merit, and susceptibility/resistance to BRD by GWAS. However, our understanding of which genetic variants drive the phenotypic variation and how genetic variants lead to phenotypic variation is still limited. The metabolome and transcriptome are important intermediate phenotypes associated with these traits, which participate in the biological processes related to phenotypic variation and are also affected by genetics. Therefore, incorporating omics data into genetic research could improve the understanding of genetic and molecular architecture of complex traits.

This thesis applied a preliminary multi-omics approach aimed to identify and refine functional SNPs and genes associated with traits of interest, and to gain insight into molecular mechanisms of how SNPs or genes affect phenotypes. In this thesis, the association between different omics levels were determined, which reveals several interesting SNPs/genes, gene networks, and molecular biomarkers. In addition, comparing our results with those from previous GWAS studies, this thesis highlights the advantages of using multiple omics data in causal SNP/gene identification, GWAS interpretation, and identification of biological mechanisms compared with single omics research.

Given that the accuracy of genomic prediction for complex traits of beef cattle still requires improvement, identifying functional SNPs and genes or biological information associated with these traits and incorporating them into genomic prediction models or evaluation panels is expected to improve the performance of genomic prediction. We recommend utilizing the functional SNPs and genes identified in our studies to perform genomic prediction and test this

hypothesis. Additionally, more research on omics data from different organs/tissues and time points using advanced data integration approaches is needed to further investigate the precise molecular mechanisms. It is anticipated that some of these elements will be generated by the BovReg project (https://www.bovreg.eu/).

In general, the results of this thesis have shed the light on improving the understanding of the genetic and molecular architecture of feed efficiency, carcass merit traits and BRD susceptibility/resistance. The outcomes of this thesis may contribute to the genetic improvement of these important traits of beef cattle, thus, further increase the profitability and sustainability of the beef industry.

6.4 References

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle - An independent validation study. *J. Anim. Sci.* 92, 2896–2904. doi:10.2527/jas.2013-7374.

Alexandre, P. A., Kogelman, L. J. A., Santana, M. H. A., Passarelli, D., Pulz, L. H., Fantinato-Neto, P., et al. (2015). Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16, 1073. doi:10.1186/s12864-015-2292-8.

Andersson, L., Archibald, A. L., Bottema, C. D., Brauning, R., Burgess, S. C., Burt, D. W., et al. (2015). Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol.* 16, 1–6. doi:10.1186/S13059-015-0622-4.

Bolormaa, S., Pryce, J. E., Kemper, K., Savin, K., Hayes, B. J., Barendse, W., et al. (2013). Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in Bos taurus, Bos indicus, and composite beef cattle. *J. Anim. Sci.* 91, 3088–3104. doi:10.2527/JAS.2012-5827.

Brandebourg, T., Wolfe, D., and Foradori, C. (2013). U.S. Beef Industry: A Sustainable Success Story, Challenges and Priorities. *J. Fish. Livest. Prod.* 01, 102. doi:10.4172/2332-2608.1000102.

Carmelo, V. A. O., Banerjee, P., da Silva Diniz, W. J., and Kadarmideen, H. N. (2020). Metabolomic networks and pathways associated with feed efficiency and related-traits in Duroc and Landrace pigs. *Sci. Rep.* 10, 255. doi:10.1038/s41598-019-57182-4.

Clemmons, B. A., Mihelic, R. I., Beckford, R. C., Powers, J. B., Melchior, E. A., McFarlane, Z. D., et al. (2017). Serum metabolites associated with feed efficiency in black angus steers. *Metabolomics* 13, 1–8. doi:10.1007/s11306-017-1282-z.

Connolly, S., Dona, A., Wilkinson-White, L., Hamblin, D., D'Occhio, M., and González, L. A. (2019). Relationship of the blood metabolome to subsequent carcass traits at slaughter in feedlot Wagyu crossbred steers. *Sci. Rep.* 9, 15139. doi:10.1038/s41598-019-51655-2.

Cooper, D. N. (2010). Functional intronic polymorphisms: Buried treasure awaiting discovery within our genes. *Hum. Genomics* 4, 284. doi:10.1186/1479-7364-4-5-284.

Deng, N., Zhou, H., Fan, H., Yuan, Y., Deng, N., Zhou, H., et al. (2017). Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget* 8, 110635–110649. doi:10.18632/ONCOTARGET.22372.

Dirks, R. A. M., Stunnenberg, H. G., and Marks, H. (2016). Genome-wide epigenomic profiling for biomarker discovery. *Clin. Epigenetics* 8, 122. doi:10.1186/S13148-016-0284-4.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017a). Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle by mapping variants to hepatic transcriptomic regions responsive to intra-mammary infection. *Genet. Sel. Evol.* 49, 44. doi:10.1186/S12711-017-0319-0.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017b). Use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits within and between dairy cattle breeds. *BMC Genomics* 18, 604. doi:10.1186/S12864-017-4004-Z.

Foroutan, A., Fitzsimmons, C., Mandal, R., Berjanskii, M. V., and Wishart, D. S. (2020). Serum metabolite biomarkers for predicting residual feed intake (RFI) of young angus bulls. *Metabolites* 10, 491. doi:10.3390/metabo10120491.

Freetly, H. C., Dickey, A., Lindholm-Perry, A. K., Thallman, R. M., Keele, J. W., Foote, A. P., et al. (2020). Digestive tract microbiota of beef cattle that differed in feed efficiency. *J. Anim. Sci.* 98, 1–16. doi:10.1093/JAS/SKAA008.

Gallois, A., Mefford, J., Ko, A., Vaysse, A., Julienne, H., Ala-Korpela, M., et al. (2019). A comprehensive study of metabolite genetics reveals strong pleiotropy and heterogeneity across time and context. *Nat. Commun.* 10, 4788. doi:10.1038/s41467-019-12703-7.

Gebreyesus, G., Bovenhuis, H., Lund, M. S., Poulsen, N. A., Sun, D., and Buitenhuis, B. (2019). Reliability of genomic prediction for milk fatty acid composition by using a multi-population reference and incorporating GWAS results. *Genet. Sel. Evol.* 51, 16. doi:10.1186/S12711-019-0460-Z.

Hu, Y., Tan, L.-J., Chen, X.-D., Liu, Z., Min, S.-S., Zeng, Q., et al. (2018). Identification of novel potentially pleiotropic variants associated with osteoporosis and opbesity using the cFDR method. *J. Clin. Endocrinol. Metab.* 103, 125–138. doi:10.1210/jc.2017-01531.

Huang, S., Chaudhary, K., and Garmire, L. X. (2017). More Is Better: Recent Progress in Multi-Omics Data Integration Methods. *Front. Genet.* 8, 84. doi:10.3389/FGENE.2017.00084.

Jiminez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front. Genet.* 12, 257. doi:10.3389/fgene.2021.627623.

Johnston, D., Earley, B., McCabe, M. S., Kim, J. W., Taylor, J. F., Lemon, K., et al. (2021). Messenger RNA biomarkers of Bovine Respiratory Syncytial Virus infection in the whole blood of dairy calves. *Sci. Rep.* 11, 9302. doi:10.1038/s41598-021-88878-1.

Johnston, D., Earley, B., McCabe, M. S., Lemon, K., Duffy, C., McMenamy, M., et al. (2019). Experimental challenge with bovine respiratory syncytial virus in dairy calves: bronchial lymph node transcriptome response. *Sci. Reports 2019 91* 9, 14736. doi:10.1038/s41598-019-51094-z.

Jorge-Smeding, E., Renand, G., Centeno, D., Pétéra, M., Durand, S., Polakof, S., et al. (2019). Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. in *EAAP Scientific Series* (Wageningen, Netherlands: Wageningen Academic Publishers), 231–232. doi:10.3920/978-90-8686-891-9_50.

Jowett, J. B. M., Curran, J. E., Johnson, M. P., Carless, M. A., Göring, H. H. H., Dyer, T. D., et al. (2010). Genetic Variation at the FTO Locus Influences RBL2 Gene Expression. *Diabetes* 59, 726–732. doi:10.2337/DB09-1277.

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

traits in beef cattle. *Livest. Sci.* 165, 200–211. doi:10.1016/j.livsci.2014.03.002.

Kim, M., Oh, I., and Ahn, J. (2018). An Improved Method for Prediction of Cancer Prognosis by Network Learning. *Genes (Basel).* 9, 478. doi:10.3390/GENES9100478.

Kreimer, A., and Pe'er, I. (2013). Variants in exons and in transcription factors affect gene expression in trans. *Genome Biol.* 14, 1–13. doi:10.1186/GB-2013-14-7-R71.

Lee, M. Y., Kim, T.-K., Walters, K.-A., and Wang, K. (2019). A biological function based biomarker panel optimization process. *Sci. Rep.* 9, 7365. doi:10.1038/s41598-019-43779-2.

Lin, Y., Zhang, W., Cao, H., Li, G., and Du, W. (2020). Classifying Breast Cancer Subtypes Using Deep Neural Networks Based on Multi-Omics Data. *Genes (Basel).* 11, 888. doi:10.3390/GENES11080888.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/S13059-014-0550-8.

Lu, D., Akanno, E. C., Crowley, J. J., Schenkel, F., Li, H., De Pauw, M., et al. (2016). Accuracy of genomic predictions for feed efficiency traits of beef cattle using 50K and imputed HD genotypes. *J. Anim. Sci.* 94, 1342–1353. doi:10.2527/JAS.2015-0126.

Ma, S., Ren, J., and Fenyö, D. (2016). Breast Cancer Prognostics Using Multi-Omics Data. *AMIA Jt. Summits Transl. Sci. Proc.* 2016, 52–59.

MacLeod, I. M., Bowman, P. J., Vander Jagt, C. J., Haile-Mariam, M., Kemper, K. E., Chamberlain, A. J., et al. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* 17, 144. doi:10.1186/S12864-016-2443-6.

McKenna, C., Keogh, K., Porter, R. K., Waters, S. M., Cormican, P., and Kenny, D. A. (2021). An examination of skeletal muscle and hepatic tissue transcriptomes from beef cattle divergent for residual feed intake. *Sci. Rep.* 11, 8942. doi:10.1038/s41598-021-87842-3.

McKenna, C., Porter, R. K., Keogh, K. A., Waters, S. M., McGee, M., and Kenny, D. A. (2018). Residual feed intake phenotype and gender affect the expression of key genes of the lipogenesis pathway in subcutaneous adipose tissue of beef cattle. *J. Anim. Sci. Biotechnol.* 9, 68. doi:10.1186/s40104-018-0282-9.

Melzer, N., Wittenburg, D., and Repsilber, D. (2013). Integrating Milk Metabolite Profile Information for the Prediction of Traditional Milk Traits Based on SNP Information for Holstein Cows. *PLoS One* 8, e70256. doi:10.1371/journal.pone.0070256.

Min, S., Lee, B., and Yoon, S. (2017). Deep learning in bioinformatics. *Brief. Bioinform.* 18, 851–869. doi:10.1093/BIB/BBW068.

Mirza, B., Wang, W., Wang, J., Choi, H., Chung, N. C., and Ping, P. (2019). Machine Learning and Integrative Analysis of Biomedical Big Data. *Genes (Basel).* 10, 87. doi:10.3390/GENES10020087.

Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Grant, J. R., Mah, J., Wang, Z., et al. (2011). Associations of marker panel scores with feed intake and efficiency traits in beef cattle using preselected single nucleotide polymorphisms. *J. Anim. Sci.* 89, 3362–3371.

doi:10.2527/JAS.2010-3362.

Nalpas, N. C., Park, S. D. E., Magee, D. A., Taraktsoglou, M., Browne, J. A., Conlon, K. M., et al. (2013). Whole-transcriptome, high-throughput RNA sequence analysis of the bovine macrophage response to Mycobacterium bovis infection in vitro. *BMC Genomics* 14, 230. doi:10.1186/1471-2164-14-230/TABLES/2.

Neibergs, H. L., Neibergs, J. S., Wojtowicz, A. J., Taylor, J. F., Seabury, C. M., and Womack, J. E. (2014a). Economic benefits of using genetic sele ction to reduce the prev alence of bovine respiratory disease complex in beef feedlot cattle. in *Proceedings of the Beef Improvement Federation (BIF) Annual Meeting and Research Symposium*, 82–86.

Neibergs, H. L., Seabury, C. M., Wojtowicz, A. J., Wang, Z., Scraggs, E., Kiser, J. N., et al. (2014b). Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. *BMC Genomics* 15, 1164. doi:10.1186/1471-2164-15-1164.

Novais, F. J., Pires, P. R. L., Alexandre, P. A., Dromms, R. A., Iglesias, A. H., Ferraz, J. B. S., et al. (2019). Identification of a metabolomic signature associated with feed efficiency in beef cattle. *BMC Genomics* 20, 8. doi:10.1186/s12864-018-5406-2.

Picard, M., Scott-Boyer, M.-P., Bodein, A., Périn, O., and Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. *Comput. Struct. Biotechnol. J.* 19, 3735–3746. doi:10.1016/J.CSBJ.2021.06.030.

Ragvin, A., Moro, E., Fredman, D., Navratilova, P., Drivenes, Ø., Engström, P. G., et al. (2010). Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc. Natl. Acad. Sci.* 107, 775–780. doi:10.1073/PNAS.0911591107.

Risso, D., Ngai, J., Speed, T. P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* 32, 896–902. doi:10.1038/nbt.2931.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. doi:10.1186/GB-2010-11-3-R25.

Sarup, P., Jensen, J., Ostersen, T., Henryon, M., and Sørensen, P. (2016). Increased prediction accuracy using a genomic feature model including prior information on quantitative trait locus regions in purebred Danish Duroc pigs. *BMC Genet.* 17, 11. doi:10.1186/s12863-015-0322-9.

Schneider, M. J., Tait, R. G., Ruble, M. V., Busby, W. D., and Reecy, J. M. (2010). Evaluation of fixed sources of variation and estimation of genetic parameters for incidence of bovine respiratory disease in preweaned calves and feedlot cattle. *J. Anim. Sci.* 88, 1220–1228. doi:10.2527/jas.2008-1755.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole blood transcriptomic analysis of beef cattle at arrival identifies potential

predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. *PLoS One* 15, e0227507. doi:10.1371/JOURNAL.PONE.0227507.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2021). Comprehensive at-arrival transcriptomic analysis of post-weaned beef cattle uncovers type I interferon and antiviral mechanisms associated with bovine respiratory disease mortality. *PLoS One* 16, e0250758. doi:10.1371/JOURNAL.PONE.0250758.

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., et al. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18, 386. doi:10.1186/s12864-017-3754-y.

Sharifi-Noghabi, H., Zolotareva, O., Collins, C. C., and Ester, M. (2019). MOLI: multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* 35, i501–i509. doi:10.1093/BIOINFORMATICS/BTZ318.

Silva, R. M. O., Fragomeni, B. O., Lourenco, D. A. L., Magalhães, A. F. B., Irano, N., Carvalheiro, R., et al. (2016). Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. *J. Anim. Sci.* 94, 3613–3623. doi:10.2527/JAS.2016-0401.

Snelling, W. M., Cushman, R. A., Keele, J. W., Maltecca, C., Thomas, M. G., Fortes, M. R. S., et al. (2013). BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection. *J. Anim. Sci.* 91, 537–552. doi:10.2527/JAS.2012-5784.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2005). Influence of breed, heterozygosity, and disease incidence on estimates of variance components of respiratory disease in preweaned beef calves. *J. Anim. Sci.* 83, 1247–1261. doi:10.2527/2005.8361247x.

Song, M., Greenbaum, J., Luttrell, J., IV, Zhou, W., Wu, C., et al. (2020). A Review of Integrative Imputation for Multi-Omics Datasets. *Front. Genet.* 11, 570255. doi:10.3389/FGENE.2020.570255.

Stranger, B. E., Stahl, E. A., and Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 187, 367–383. doi:10.1534/genetics.110.120907.

Suhre, K., and Gieger, C. (2012). Genetic variation in metabolic phenotypes: study designs and applications. *Nat. Rev. Genet.* 13, 759–769. doi:10.1038/nrg3314.

Sun, H. Z., Srithayakumar, V., Jiminez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics* 112, 3968–3977. doi:10.1016/J.YGENO.2020.07.014.

Sun, Y. V., and Hu, Y. J. (2016). Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv. Genet.* 93, 147–190. doi:10.1016/BS.ADGEN.2015.11.004.

Tizioto, P. C., Kim, J. W., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015). Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: An RNA-sequence analysis of the bronchial lymph

node transcriptome. *PLoS One* 10, e0131459. doi:10.1371/journal.pone.0131459.

Wang, X., and Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites* 9, 151. doi:10.3390/metabo9070151.

Wang, X., Wang, K., Radovich, M., Wang, Y., Wang, G., Feng, W., et al. (2009). Genome-wide prediction of cis-acting RNA elements regulating tissue-specific pre-mRNA alternative splicing. *BMC Genomics* 10, S4. doi:10.1186/1471-2164-10-S1-S4.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Yan, J., Risacher, S. L., Shen, L., and Saykin, A. J. (2018). Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. *Brief. Bioinform.* 19, 1370–1381. doi:10.1093/bib/bbx066.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020a). 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, 36. doi:10.1186/s12864-019-6362-1.

Zhang, Y., Parmigiani, G., and Johnson, W. E. (2020b). ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma.* 2, lqaa078. doi:10.1093/NARGAB/LQAA078.

Zhao, S., and Li, L. (2020). Chemical derivatization in LC-MS-based metabolomics study. *TrAC Trends Anal. Chem.* 131, 115988. doi:10.1016/J.TRAC.2020.115988.

Zhou, G., Li, S., and Xia, J. (2020). "Network-Based Approaches for Multi-omics Integration," in *Computational Methods and Data Analysis for Metabolomics* (Humana, New York, NY), 469–487. doi:10.1007/978-1-0716-0239-3_23.

Bibliography

Abo-Ismail, M. K., Lansink, N., Akanno, E., Karisa, B. K., Crowley, J. J., Moore, S. S., et al. (2018). Development and validation of a small SNP panel for feed efficiency in beef cattle. *J. Anim. Sci*. 96, 375–397. doi:10.1093/jas/sky020.

Abo-Ismail, M. K., Vander Voort, G., Squires, J. J., Swanson, K. C., Mandell, I. B., Liao, X., et al. (2014). Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. *BMC Genet*. 15, 14. doi:10.1186/1471-2156-15-14.

Ackermann, M. R., Derscheid, R., and Roth, J. A. (2010). Innate Immunology of Bovine Respiratory Disease. *Vet. Clin. North Am. Food Anim. Pract*. 26, 215. doi:10.1016/J.CVFA.2010.03.001.

Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., and Misztal, I. (2019). Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol*. 51, 28. doi:10.1186/s12711-019-0469-3.

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

Ahola, J. K., and Hill, R. A. (2012). "Input Factors Affecting Profitability: A Changing Paradigm and a Challenging Time," in Feed Efficiency in the Beef Industry (Hoboken, New Jersey: Wiley-Blackwell), 7–19. doi:10.1002/9781118392331.ch1.

Akanno, E. C., Plastow, G., Woodward, B. W., Bauck, S., Okut, H., Wu, X. L., et al. (2014). Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle - An independent validation study. *J. Anim. Sci*. 92, 2896–2904. doi:10.2527/jas.2013-7374.

Akdis, M., Burgler, S., Crameri, R., Eiwegger, T., Fujita, H., Gomez, E., et al. (2011). Interleukins, from 1 to 37, and interferon-γ: Receptors, functions, and roles in diseases. *J. Allergy Clin. Immunol*. 127, 701-721.e70. doi:10.1016/j.jaci.2010.11.050.

Akram, M. (2014). Citric acid cycle and role of its intermediates in metabolism. *Cell Biochem. Biophys*. 68, 475–478. doi:10.1007/s12013-013-9750-1.

Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12, 246. doi:10.1186/1471-2105-12-246.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109.

Alexandre, P. A., Kogelman, L. J. A., Santana, M. H. A., Passarelli, D., Pulz, L. H., Fantinato-Neto, P., et al. (2015). Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16, 1073. doi:10.1186/s12864-015-2292-8.

Andersson, L., Archibald, A. L., Bottema, C. D., Brauning, R., Burgess, S. C., Burt, D. W., et al. (2015). Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol*. 16, 1–6. doi:10.1186/S13059-015-0622-4.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Aoyama, C., Liao, H., and Ishidate, K. (2004). Structure and function of choline kinase isoforms in mammalian cells. *Prog. Lipid Res*. 43, 266–281. doi:10.1016/j.plipres.2003.12.001.

Archer, J. A., Barwick, S. A., and Graser, H. U. (2004). Economic evaluation of beef cattle breeding schemes incorporating performance testing of young bulls for feed intake. *Aust. J. Exp. Agric*. 44, 393–404. doi:10.1071/ea02054.

Archer, J. A., Richardson, E. C., Herd, R. M., and Arthur, P. F. (1999). Potential for selection to improve efficiency of feed use in beef cattle: A review. Aust. *J. Agric. Res*. 50, 147–161. doi:10.1071/a98075.

Arena, W. P., Malyak, M., Guthridge, C. J., and Gabay, C. (1998). Interleukin-1 receptor antagonist: Role in biology. *Annu. Rev. Immunol*. 16, 27–55. doi:10.1146/annurev.immunol.16.1.27.

Arthur, P. F., Archer, J. A., Herd, R. M., and Melville, G. J. (2001a). Response To Selection For Net Feed Intake in Beef Cattle. in Proceedings of the Association for the Advancement of Animal Breedings and Genetics, 135–138.

Arthur, P. F., Archer, J. A., Johnston, D. J., Herd, R. M., Richardson, E. C., and Parnell, P. F. (2001b). Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *J. Anim. Sci*. 79, 2805. doi:10.2527/2001.79112805x.

Arthur, P. F., Renand, G., and Krauss, D. (2001). Genetic and phenotypic relationships among different measures of growth and feed efficiency in young Charolais bulls. *Livest. Prod. Sci*. 68, 131–139. doi:10.1016/S0301-6226(00)00243-8.

Baldassini, W. A., Bonilha, S. F. M., Branco, R. H., Vieira, J. C. S., Padilha, P. M., and Lanna, D. P. D. (2018). Proteomic investigation of liver from beef cattle (Bos indicus) divergently ranked on residual feed intake. *Mol. Biol. Rep*. 45, 2765–2773. doi:10.1007/s11033-018-4341-2.

Baptista, A. L., Rezende, A. L., Fonseca, P. de A., Massi, R. P., Nogueira, G. M., Magalhães, L. Q., et al. (2017). Bovine respiratory disease complex associated mortality and morbidity rates in feedlot cattle from southeastern Brazil. *J. Infect. Dev. Ctries*. 11, 791–799. doi:10.3855/jidc.9296.

Basarab, J. A. A., Colazo, M. G. G., Ambrose, D. J. J., Novak, S., McCartney, D., and Baron, V. S. S. (2011). Residual feed intake adjusted for backfat thickness and feeding frequency is independent of fertility in beef heifers. Can. *J. Anim. Sci*. 91, 573–584. doi:10.4141/cjas2011-010.

Basarab, J. A., Price, M. A., Aalhus, J. L., Okine, E. K., Snelling, W. M., and Lyle, K. L. (2003). Residual feed intake and body composition in young growing cattle. *Can. J. Anim. Sci*. 83, 189–204. doi:10.4141/A02-065.

Bernardo, R. (2008). Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Sci*. 48, 1649–1664. doi:10.2135/CROPSCI2008.03.0131.

Berry, D. P., and Crowley, J. J. (2012). Residual intake and body weight gain: A new measure of efficiency in growing cattle. *J. Anim. Sci*. 90, 109–115. doi:10.2527/JAS.2011-4245.

Beuchel, C., Becker, S., Dittrich, J., Kirsten, H., Toenjes, A., Stumvoll, M., et al. (2019). Clinical and lifestyle related factors influencing whole blood metabolite levels – A comparative analysis of three large cohorts. *Mol. Metab*. 29, 76–85. doi:10.1016/j.molmet.2019.08.010.

Blakebrough-Hall, C., Dona, A., D'occhio, M. J., McMeniman, J., and González, L. A. (2020). Diagnosis of Bovine Respiratory Disease in feedlot cattle using blood 1H NMR metabolomics. *Sci. Rep*. 10, 115. doi:10.1038/s41598-019-56809-w.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.

Bolormaa, S., Hayes, B. J., Savin, K., Hawken, R., Barendse, W., Arthur, P. F., et al. (2011). Genome-wide association studies for feedlot and growth traits in cattle. *J. Anim. Sci*. 89, 1684–1697. doi:10.2527/jas.2010-3079.

Bolormaa, S., Pryce, J. E., Kemper, K., Savin, K., Hayes, B. J., Barendse, W., et al. (2013). Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in Bos taurus, Bos indicus, and composite beef cattle. *J. Anim. Sci*. 91, 3088–3104. doi:10.2527/JAS.2012-5827.

Bowland, S. L., and Shewen, P. E. (2000). Bovine respiratory disease: Commercial vaccines currently available in Canada. *Can. Vet*. J. 41, 33–48.

Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169, 1177. doi:10.1016/J.CELL.2017.05.038.

Brandebourg, T., Wolfe, D., and Foradori, C. (2013). U.S. Beef Industry: A Sustainable Success Story, Challenges and Priorities. *J. Fish. Livest. Prod*. 01, 102. doi:10.4172/2332-2608.1000102.

Brody, S. (1945). Bioenergetics and growth; with special reference to the efficiency complex in domestic animals. New York, USA: Reinhold Publishing Corp Available at: https://psycnet.apa.org/record/1946-03094-000.

Brouard, J. S., Schenkel, F., Marete, A., and Bissonnette, N. (2019). The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J. Anim. Sci. Biotechnol*. 10, 44. doi:10.1186/s40104-019-0359-0.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet*. 103, 338–348. doi:10.1016/j.ajhg.2018.07.015.

Buitenhuis, A. J. J., Sundekilde, U. K. K., Poulsen, N. A. A., Bertram, H. C. C., Larsen, L. B. B., and Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait

loci for metabolites in Danish Holstein milk. *J. Dairy Sci*. 96, 3285–3295. doi:10.3168/jds.2012-5914.

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678. doi:10.1038/nature05911.

Calvano, S. E., Xiao, W., Richards, D. R., Felciano, R. M., Baker, H. V., Cho, R. J., et al. (2005). A network-based analysis of systemic inflammation in humans. *Nature* 437, 1032–1037. doi:10.1038/nature03985.

Cánovas, A., Reverter, A., DeAtley, K. L., Ashley, R. L., Colgrave, M. L., Fortes, M. R. S., et al. (2014). Multi-Tissue Omics Analyses Reveal Molecular Regulatory Networks for Puberty in Composite Beef Cattle. *PLoS One* 9, e102551. doi:10.1371/JOURNAL.PONE.0102551.

Cantalapiedra-Hijar, G., Abo-Ismail, M., Carstens, G. E., Guan, L. L., Hegarty, R., Kenny, D. A., et al. (2018). Review: Biological determinants of between-animal variation in feed efficiency of growing beef cattle. *Animal* 12, S321–S335. doi:10.1017/S1751731118001489.

Carmelo, V. A. O., Banerjee, P., da Silva Diniz, W. J., and Kadarmideen, H. N. (2020). Metabolomic networks and pathways associated with feed efficiency and related-traits in Duroc and Landrace pigs. *Sci. Rep*. 10, 255. doi:10.1038/s41598-019-57182-4.

Carstens, G., and Tedeschi, L. (2006). Defining feed efficiency in beef cattle. in Proceedings of Beef Improvement Federation 38th Annual Research Symposium and Annual Meeting, 12–21.

Casas, E., Garcia, M. D., Wells, J. E., and Smith, T. P. L. (2011). Association of single nucleotide polymorphisms in the ANKRA2 and CD180 genes with bovine respiratory disease and presence of Mycobacterium avium subsp. paratuberculosis1. *Anim. Genet*. 42, 571–577. doi:10.1111/j.1365-2052.2011.02189.x.

Ceacero, T. M., Mercadante, M. E. Z., Cyrillo, J. N. D. S. G., Canesin, R. C., Bonilha, S. F. M., and De Albuquerque, L. G. (2016). Phenotypic and genetic correlations of feed efficiency traits with growth and carcass traits in nellore cattle selected for postweaning weight. *PLoS One* 11, e0161366. doi:10.1371/journal.pone.0161366.

Chamorro, M. F., and Palomares, R. A. (2020). Bovine Respiratory Disease Vaccination Against Viral Pathogens: Modified-Live Versus Inactivated Antigen Vaccines, Intranasal Versus Parenteral, What Is the Evidence? *Vet. Clin. North Am. Food Anim. Pract*. 36, 461. doi:10.1016/J.CVFA.2020.03.006.

Chan, E. K. F., Rowe, H. C., Hansen, B. G., and Kliebenstein, D. J. (2010). The Complex Genetic Architecture of the Metabolome. *PLoS Genet*. 6, e1001198. doi:10.1371/journal.pgen.1001198.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. doi:10.1186/S13742-015-0047-8.

Chang, T., Xia, J., Xu, L., Wang, X., Zhu, B., Zhang, L., et al. (2018). A genome-wide association study suggests several novel candidate genes for carcass traits in Chinese Simmental beef cattle. *Anim. Genet*. 49, 312–316. doi:10.1111/AGE.12667.

Chang, W., Hatch, G. M., Wang, Y., Yu, F., and Wang, M. (2019). The relationship between phospholipids and insulin resistance: From clinical to experimental studies. *J. Cell. Mol. Med*. 23, 702–710. doi:10.1111/jcmm.13984.

Chen, L., Mao, F., Crews, D. H., Vinsky, M., and Li, C. (2014). Phenotypic and genetic relationships of feeding behavior with feed intake, growth performance, feed efficiency, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci*. 92, 974–983. doi:10.2527/jas.2013-6926.

Chen, L., Schenkel, F., Vinsky, M., Crews, D. H., and Li, C. (2013). Accuracy of predicting genomic breeding values for residual feed intake in Angus and Charolais beef cattle. *J. Anim. Sci*. 91, 4669–4678. doi:10.2527/JAS.2013-5715.

Chen, L., Vinsky, M., and Li, C. (2015). Accuracy of predicting genomic breeding values for carcass merit traits in Angus and Charolais beef cattle. *Anim. Genet*. 46, 55–59. doi:10.1111/age.12238.

Chen, Y., Gondro, C., Quinn, K., Herd, R. M., Parnell, P. F., and Vanselow, B. (2011). Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim. Genet*. 42, 475–490. doi:10.1111/j.1365-2052.2011.02182.x.

Chi, J., VanLeeuwen, J. A., Weersink, A., and Keefe, G. P. (2002). Direct production losses and treatment costs from bovine viral diarrhoea virus, bovine leukosis virus, Mycobacterium avium subspecies paratuberculosis, and Neospora caninum. *Prev. Vet. Med*. 55, 137–153. doi:10.1016/S0167-5877(02)00094-6.

Chirase, N. K., and Greene, L. W. (2001). Dietary zinc and manganese sources administered from the fetal stage onwards affect immune response of transit stressed and virus infected offspring steer calves. *Anim. Feed Sci. Technol*. 93, 217–228. doi:10.1016/S0377-8401(01)00277-2.

Clemmons, B. A., Mihelic, R. I., Beckford, R. C., Powers, J. B., Melchior, E. A., McFarlane, Z. D., et al. (2017). Serum metabolites associated with feed efficiency in black angus steers. *Metabolomics* 13, 1–8. doi:10.1007/s11306-017-1282-z.

Clish, C. B. (2015). Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb. Mol. Case Stud*. 1, a000588. doi:10.1101/MCS.A000588.

Connolly, S., Dona, A., Hamblin, D., D'Occhio, M. J., and González, L. A. (2020). Changes in the blood metabolome of Wagyu crossbred steers with time in the feedlot and relationships with marbling. *Sci. Rep*. 10, 18987. doi:10.1038/s41598-020-76101-6.

Connolly, S., Dona, A., Wilkinson-White, L., Hamblin, D., D'Occhio, M., and González, L. A. (2019). Relationship of the blood metabolome to subsequent carcass traits at slaughter in feedlot Wagyu crossbred steers. *Sci. Rep*. 9, 15139. doi:10.1038/s41598-019-51655-2.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. *Nat. Rev. Genet*. 10, 184–194. doi:10.1038/nrg2537.

Cooper, D. N. (2010). Functional intronic polymorphisms: Buried treasure awaiting discovery within our genes. *Hum. Genomics* 4, 284. doi:10.1186/1479-7364-4-5-284.

Cooper, G. M., and Hausman, R. E. (2000). The Cell: A Molecular Approach. Washington, DC: ASM press.

Costa-Silva, J., Domingues, D., and Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* 12, e0190152. doi:10.1371/JOURNAL.PONE.0190152.

Dabo, S. M., Taylor, J. D., and Confer, A. W. (2008). Pasteurella multocida and bovine respiratory disease. *Anim. Heal. Res. Rev*. 8, 129–150. doi:10.1017/S1466252307001399.

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., Van Binsbergen, R., Brøndum, R. F., et al. (2014). Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet*. 46, 858–865. doi:10.1038/ng.3034.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/BIOINFORMATICS/BTR330.

De Roos, A. P. W., Hayes, B. J., and Goddard, M. E. (2009). Reliability of genomic predictions across multiple populations. *Genetics* 183, 1545–1553. doi:10.1534/genetics.109.104935.

Decaro, N., Campolo, M., Desario, C., Cirone, F., D'abramo, M., Lorusso, E., et al. (2008). Respiratory Disease Associated with Bovine Coronavirus Infection in Cattle Herds in Southern Italy. *J. Vet. Diagnostic Investig*. 20, 28–32. doi:10.1177/104063870802000105.

Demirkan, A., van Duijn, C. M., Ugocsai, P., Isaacs, A., Pramstaller, P. P., Liebisch, G., et al. (2012). Genome-wide association study identifies novel loci associated with circulating phospho- and sphingolipid concentrations. *PLoS Genet*. 8, e1002490. doi:10.1371/journal.pgen.1002490.

Deng, N., Zhou, H., Fan, H., Yuan, Y., Deng, N., Zhou, H., et al. (2017). Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget* 8, 110635–110649. doi:10.18632/ONCOTARGET.22372.

Dinarello, C. A. (2018). Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol. Rev*. 281, 8. doi:10.1111/IMR.12621.

Dirks, R. A. M., Stunnenberg, H. G., and Marks, H. (2016). Genome-wide epigenomic profiling for biomarker discovery. *Clin. Epigenetics* 8, 122. doi:10.1186/S13148-016-0284-4.

Do, D. N., Schenkel, F. S., Miglior, F., Zhao, X., and Ibeagha-Awemu, E. M. (2018). Genome wide association study identifies novel potential candidate genes for bovine milk cholesterol content. *Sci. Rep*. 8, 13239. doi:10.1038/s41598-018-31427-0.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.

Doublet, A. C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., et al. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genet. Sel. Evol*. 51, 52. doi:10.1186/S12711-019-0495-1/FIGURES/7.

Druhan, L. J., Lance, A., Li, S., Price, A. E., Emerson, J. T., Baxter, S. A., et al. (2017). Leucine Rich α-2 Glycoprotein: A Novel Neutrophil Granule Protein and Modulator of Myelopoiesis. *PLoS One* 12, e0170261. doi:10.1371/JOURNAL.PONE.0170261.

Duncan, L., Yilmaz, Z., Walters, R., Goldstein, J., Anttila, V., Bulik-Sullivan, B., et al. (2017). Genome-Wide Association Study Reveals First Locus for Anorexia Nervosa and Metabolic Correlations. *Am. J. Psychiatry* 174, 850. doi:10.1176/APPI.AJP.2017.16121402.

Edgeworth, J., Gorman, M., Bennett, R., Freemont, P., and Hogg, N. (1991). Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J. Biol. Chem*. 266, 7706–7713. doi:10.1016/S0021-9258(20)89506-4.

Ekine-Dzivenu, C., Vinsky, M., Basarab, J. A., Aalhus, J. L., Dugan, M. E. R., and Li, C. (2017). Phenotypic and genetic correlations of fatty acid composition in subcutaneous adipose tissue with carcass merit and meat tenderness traits in Canadian beef cattle. *J. Anim. Sci*. 95, 5184–5196. doi:10.2527/jas2017.1966.

Emam, M., Livernois, A., Paibomesai, M., Atalla, H., and Mallard, B. (2019). Genetic and Epigenetic Regulation of Immune Response and Resistance to Infectious Diseases in Domestic Ruminants. *Vet. Clin. North Am. - Food Anim. Pract*. 35, 405–429. doi:10.1016/j.cvfa.2019.07.002.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017). Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle by mapping variants to hepatic transcriptomic regions responsive to intra-mammary infection. *Genet. Sel. Evol*. 49, 44. doi:10.1186/S12711-017-0319-0.

Fang, L., Sahana, G., Ma, P., Su, G., Yu, Y., Zhang, S., et al. (2017b). Use of biological priors enhances understanding of genetic architecture and genomic prediction of complex traits within and between dairy cattle breeds. *BMC Genomics* 18, 604. doi:10.1186/S12864-017-4004-Z.

Fernandes Júnior, G. A., Rosa, G. J. M., Valente, B. D., Carvalheiro, R., Baldi, F., Garcia, D. A., et al. (2016). Genomic prediction of breeding values for carcass traits in Nellore cattle. *Genet. Sel. Evol*. 48, 7. doi:10.1186/s12711-016-0188-y.

Ferrero-Miliani, L., Nielsen, O. H., Andersen, P. S., and Girardin, S. E. (2007). Chronic inflammation: Importance of NOD2 and NALP3 in interleukin-1β generation. *Clin. Exp. Immunol*. 147, 227–235. doi:10.1111/j.1365-2249.2006.03261.x.

Flint, J., Timpson, N., and Munafò, M. (2014). Assessing the utility of intermediate phenotypes for genetic mapping of psychiatric disease. *Trends Neurosci*. 37, 733. doi:10.1016/J.TINS.2014.08.007.

Fonseca, L. D., Eler, J. P., Pereira, M. A., Rosa, A. F., Alexandre, P. A., Moncau, C. T., et al. (2019). Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. *Sci. Rep*. 9, 5364. doi:10.1038/s41598-019-41813-x.

Fonseca, P. A. de S., Id-Lahoucine, S., Reverter, A., Medrano, J. F., Fortes, M. S., Casellas, J., et al. (2018). Combining multi-OMICs information to identify key-regulator genes for pleiotropic effect on fertility and production traits in beef cattle. *PLoS One* 13, e0205295. doi:10.1371/journal.pone.0205295.

Fontanesi, L. (2016). “Merging Metabolomics, Genetics, and Genomics in Livestock to Dissect Complex Production Traits,” in Systems Biology in Animal Production and Health, Vol. 1 (Springer, Cham), 43–62. doi:10.1007/978-3-319-43335-6_3.

Fontanesi, L. (2016). Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. *Anim. Front*. 6, 73. doi:10.2527/af.2016-0011.

Foroutan, A., Fitzsimmons, C., Mandal, R., Berjanskii, M. V., and Wishart, D. S. (2020). Serum metabolite biomarkers for predicting residual feed intake (RFI) of young angus bulls. *Metabolites* 10, 491. doi:10.3390/metabo10120491.

Freetly, H. C., Dickey, A., Lindholm-Perry, A. K., Thallman, R. M., Keele, J. W., Foote, A. P., et al. (2020). Digestive tract microbiota of beef cattle that differed in feed efficiency. *J. Anim. Sci*. 98, 1–16. doi:10.1093/JAS/SKAA008.

Frendl, G. (1992). Interleukin 3: From colony-stimulating factor to pluripotent immunoregulatory cytokine. *Int. J. Immunopharmacol*. 14, 421–430. doi:10.1016/0192-0561(92)90172-H.

Frost, H. R., and Amos, C. I. (2018). A multi-omics approach for identifying important pathways and genes in human cancer. *BMC Bioinformatics* 19, 479. doi:10.1186/S12859-018-2476-8.

Fu, Y., Wang, L., Tang, Z., Yin, D., Xu, J., Fan, Y., et al. (2020). An integration analysis based on genomic, transcriptomic and QTX information reveals credible candidate genes for fat-related traits in pigs. *Anim. Genet*. 51, 683–693. doi:10.1111/AGE.12971.

Fujimoto, M., Matsumoto, T., Serada, S., Tsujimura, Y., Hashimoto, S., Yasutomi, Y., et al. (2020). Leucine-rich alpha 2 glycoprotein is a new marker for active disease of tuberculosis. *Sci. Rep*. 10, 3384. doi:10.1038/s41598-020-60450-3.

Gallois, A., Mefford, J., Ko, A., Vaysse, A., Julienne, H., Ala-Korpela, M., et al. (2019). A comprehensive study of metabolite genetics reveals strong pleiotropy and heterogeneity across time and context. *Nat. Commun*. 10, 4788. doi:10.1038/s41467-019-12703-7.

García-Ruiz, A., Cole, J. B., VanRaden, P. M., Wiggans, G. R., Ruiz-López, F. J., and Tassell, C. P. Van (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci*. 113, E3995–E4004. doi:10.1073/PNAS.1519061113.

Garcia, M. D., Thallman, R. M., Wheeler, T. L., Shackelford, S. D., and Casas, E. (2010). Effect of bovine respiratory disease and overall pathogenic disease incidence on carcass traits. *J. Anim. Sci*. 88, 491–496. doi:10.2527/jas.2009-1874.

Gebreyesus, G., Bovenhuis, H., Lund, M. S., Poulsen, N. A., Sun, D., and Buitenhuis, B. (2019). Reliability of genomic prediction for milk fatty acid composition by using a multi-population reference and incorporating GWAS results. *Genet. Sel. Evol*. 51, 16. doi:10.1186/S12711-019-0460-Z.

Georges, M., Charlier, C., and Hayes, B. (2018). Harnessing genomic information for livestock improvement. *Nat. Rev. Genet*. 20, 135–156. doi:10.1038/s41576-018-0082-2.

Gerber, P. J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., et al. (2013). Tackling climate change through livestock – A global assessment of emissions and

mitigation opportunities. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).

Gieger, C., Geistlinger, L., Altmaier, E., Hrabé de Angelis, M., Kronenberg, F., Meitinger, T., et al. (2008). Genetics meets metabolomics: A genome-wide association study of metabolite profiles in human serum. *PLoS Genet*. 4, 1000282. doi:10.1371/journal.pgen.1000282.

Gilmour, A. R., Gogel, B. J., and Welham, S. J. (2015). ASReml User Guide Functional Specification. Hemel Hempstead: VSN International Ltd.

Goddard, M. (2009). Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica* 136, 245–257. doi:10.1007/s10709-008-9308-0.

Goddard, M. E., and Hayes, B. J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet*. 10, 381–391. doi:10.1038/nrg2575.

Goldansaz, S. A., Markus, S., Berjanskii, M., Rout, M., Guo, A. C., Wang, Z., et al. (2020). Candidate serum metabolite biomarkers of residual feed intake and carcass merit in sheep. *J. Anim. Sci*. 98, 1–15. doi:10.1093/JAS/SKAA298.

Griffin, D. (1997). Economic impact associated with respiratory disease in beef cattle. *Vet. Clin. North Am. Food Anim. Pract*. 13, 367–377. doi:10.1016/S0749-0720(15)30302-9.

Griffin, D., Chengappa, M. M., Kuszak, J., and McVey, D. S. (2010). Bacterial pathogens of the bovine respiratory disease complex. *Vet. Clin. North Am. - Food Anim. Pract*. 26, 381–394. doi:10.1016/j.cvfa.2010.04.004.

Grigoletto, L., Ferraz, J. B. S., Oliveira, H. R., Eler, J. P., Bussiman, F. O., Abreu Silva, B. C., et al. (2020). Genetic Architecture of Carcass and Meat Quality Traits in Montana Tropical® Composite Beef Cattle. *Front. Genet*. 0, 123. doi:10.3389/FGENE.2020.00123.

Haas, R., Zelezniak, A., Iacovacci, J., Kamrad, S., Townsend, S. J., and Ralser, M. (2017). Designing and interpreting "multi-omic" experiments that may change our understanding of biology. *Curr. Opin. Syst. Biol*. 6, 37–45. doi:10.1016/j.coisb.2017.08.009.

Habier, D., Fernando, R. L., and Dekkers, J. C. M. (2007). The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics* 177, 2389–2397. doi:10.1534/GENETICS.107.081190.

Hanset, R., and Michaux, C. (1982). "Creatine and Creatinine Levels in Plasma, Red Cells and Muscles as Characteristics of Double-Muscled Cattle," in Muscle Hypertrophy of Genetic Origin and its use to Improve Beef Production (Springer Netherlands), 237–256. doi:10.1007/978-94-009-7550-7_21.

Hasin, Y., Seldin, M., and Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biol*. 18, 1–15. doi:10.1186/s13059-017-1215-1.

Hay, E. H., and Roberts, A. (2018). Genome-wide association study for carcass traits in a composite beef cattle breed. *Livest. Sci*. 213, 35–43. doi:10.1016/J.LIVSCI.2018.04.018.

Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K., and Goddard, M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol*. 41, 51. doi:10.1186/1297-9686-41-51/FIGURES/2.

Hayes, B. J., Visscher, P. M., and Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res. (Camb)*. 91, 47–60. doi:10.1017/S0016672308009981.

Hegarty, R. S., Goopy, J. P., Herd, R. M., and McCorkell, B. (2007). Cattle selected for lower residual feed intake have reduced daily methane production. *J. Anim. Sci*. 85, 1479–1486. doi:10.2527/jas.2006-236.

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., and Edwards, T. L. (2017). Population stratification in genetic association studies. *Curr. Protoc. Hum. Genet*. 95, 1.22.1-1.22.23. doi:10.1002/cphg.48.

Herd, R. M., and Bishop, S. C. (2000). Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. *Livest. Prod. Sci*. 63, 111–119. doi:10.1016/S0301-6226(99)00122-0.

Higgins, M. G., Fitzsimons, C., McClure, M. C., McKenna, C., Conroy, S., Kenny, D. A., et al. (2018). GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle. *Sci. Rep*. 8, 14301. doi:10.1038/s41598-018-32374-6.

Higgins, M. G., Kenny, D. A., Fitzsimons, C., Blackshields, G., Coyle, S., McKenna, C., et al. (2019). The effect of breed and diet type on the global transcriptome of hepatic tissue in beef cattle divergent for feed efficiency. *BMC Genomics* 20, 525. doi:10.1186/s12864-019-5906-8.

Hoff, J. L., Decker, J. E., Schnabel, R. D., Seabury, C. M., Neibergs, H. L., and Taylor, J. F. (2019). QTL-mapping and genomic prediction for bovine respiratory disease in U.S. Holsteins using sequence imputation and feature selection. *BMC Genomics* 20, 555. doi:10.1186/s12864-019-5941-5.

Holland, R., and Loveday, D. (2013). Understanding Yield Grades and Quality Grades for Value-Added Beef Producers and Marketers. Institute of Agriculture, University of Tennessee, Knoxville.

Hosaka, K., Tanaka, S., Nikawa, J. ichi, and Yamashita, S. (1992). Cloning of a human choline kinase cDNA by complementation of the yeast cki mutation. *FEBS Lett*. 304, 229–232. doi:10.1016/0014-5793(92)80625-Q.

Hrdlickova, R., Toloue, M., and Tian, B. (2017). RNA-Seq methods for transcriptome analysis. Wiley Interdiscip. *Rev. RNA* 8, e1364. doi:10.1002/WRNA.1364.

Hu, Y., Tan, L.-J., Chen, X.-D., Liu, Z., Min, S.-S., Zeng, Q., et al. (2018). Identification of novel potentially pleiotropic variants associated with osteoporosis and opbesity using the cFDR method. *J. Clin. Endocrinol. Metab*. 103, 125–138. doi:10.1210/jc.2017-01531.

Huang, Q. C., Xu, Z. R., Han, X. Y., and Li, W. F. (2008). Effect of dietary betaine supplementation on lipogenic enzyme activities and fatty acid synthase mRNA expression in finishing pigs. *Anim. Feed Sci. Technol*. 140, 365–375. doi:10.1016/j.anifeedsci.2007.03.007.

Huang, S., Chaudhary, K., and Garmire, L. X. (2017). More Is Better: Recent Progress in Multi-Omics Data Integration Methods. *Front. Genet*. 8, 84. doi:10.3389/FGENE.2017.00084.

Hwang, Y. H., and Joo, S. T. (2016). Fatty acid profiles of ten muscles from high and low marbled (Quality Grade 1++ and 2) hanwoo steers. *Korean J. Food Sci. Anim. Resour*. 36, 679–688. doi:10.5851/kosfa.2016.36.5.679.

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., et al. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet*. 48, 1031–1036. doi:10.1038/NG.3623.

Illig, T., Gieger, C., Zhai, G., Römisch-Margl, W., Wang-Sattler, R., Prehn, C., et al. (2010). A genome-wide perspective of genetic variation in human metabolism. *Nat. Genet*. 42, 137–141. doi:10.1038/ng.507.

Ioannidis, J. P. A. (2007). Non-Replication and Inconsistency in the Genome-Wide Association Setting. *Hum. Hered*. 64, 203–213. doi:10.1159/000103512.

Irsik, M., Langemeier, M., Schroeder, T., Spire, M., and Deen Roder, J. (2006). Estimating the Effects of Animal Health on the Performance of Feedlot Cattle. *Bov. Pract*. 40, 65–74. doi:org/10.21423/bovine-vol40no2p65-74.

Ives, S. E., and Richeson, J. T. (2015). Use of Antimicrobial Metaphylaxis for the Control of Bovine Respiratory Disease in High-Risk Cattle. *Vet. Clin. North Am. - Food Anim. Pract*. 31, 341–350. doi:10.1016/j.cvfa.2015.05.008.

Jeong, J. Y., Kim, M., Ji, S. Y., Baek, Y. C., Lee, S., Oh, Y. K., et al. (2020). Metabolomics analysis of the beef samples with different meat qualities and tastes. *Food Sci. Anim. Resour*. 40, 924–937. doi:10.5851/KOSFA.2020.E59.

Jericho, K. W. F., and Langford, E. V. (1978). Pneumonia in Calves Produced with Aerosols of Bovine Herpesvirus 1 and Pasteurella haemolytica. *Can. J. Comp. Med*. 42, 269.

Jiang, T., Huang, Z., Zhang, S., Zou, W., Xiang, L., Wu, X., et al. (2018). miR-23b inhibits proliferation of SMMC-7721 cells by directly targeting IL-11. *Mol. Med. Rep*. 18, 1591–1599. doi:10.3892/MMR.2018.9151.

Jiminez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front. Genet*. 12, 257. doi:10.3389/fgene.2021.627623.

Johnston, D., Earley, B., McCabe, M. S., Kim, J. W., Taylor, J. F., Lemon, K., et al. (2021). Messenger RNA biomarkers of Bovine Respiratory Syncytial Virus infection in the whole blood of dairy calves. *Sci. Rep*. 11, 9302. doi:10.1038/s41598-021-88878-1.

Johnston, D., Earley, B., McCabe, M. S., Lemon, K., Duffy, C., McMenamy, M., et al. (2019). Experimental challenge with bovine respiratory syncytial virus in dairy calves: bronchial lymph node transcriptome response. *Sci. Reports* 2019 91 9, 14736. doi:10.1038/s41598-019-51094-z.

Jorge-Smeding, E., Renand, G., Centeno, D., Pétéra, M., Durand, S., Polakof, S., et al. (2019). Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. in EAAP Scientific Series (Wageningen, Netherlands: Wageningen Academic Publishers), 231–232. doi:10.3920/978-90-8686-891-9_50.

Jowett, J. B. M., Curran, J. E., Johnson, M. P., Carless, M. A., Göring, H. H. H., Dyer, T. D., et al. (2010). Genetic Variation at the FTO Locus Influences RBL2 Gene Expression. *Diabetes* 59, 726–732. doi:10.2337/DB09-1277.

Karisa, B. K., Thomson, J., Wang, Z., Li, C., Montanholi, Y. R., Miller, S. P., et al. (2014). Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livest. Sci.* 165, 200–211. doi:10.1016/j.livsci.2014.03.002.

Kazala, E. C., Lozeman, F. J., Mir, P. S., Aalhus, J. L., Schmutz, S. M., and Weselake, R. J. (2006). Fatty acid composition of muscle fat and enzymes of storage lipid synthesis in whole muscle from beef cattle. *Lipids* 41, 1049–1057. doi:10.1007/s11745-006-5055-0.

Kazala, E. C., Lozeman, F. J., Mir, P. S., Laroche, A., Bailey, D. R. C., and Weselake, R. J. (1999). Relationship of fatty acid composition to intramuscular fat content in beef from crossbred wagyu cattle. *J. Anim. Sci*. 77, 1717–1725. doi:10.2527/1999.7771717x.

Kellner, O. J., and Goodwin, W. (1913). The Scientific Feeding of Animals. London, UK: Duckworth & Co. Available at: https://books.google.ca/books?id=iv01AQAAIAAJ&dq=The+Scientific+Feeding+of+Animals&lr=&source=gbs_navlinks_s.

Kettunen, J., Tukiainen, T., Sarin, A. P., Ortega-Alonso, A., Tikkanen, E., Lyytikäinen, L. P., et al. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet*. 44, 269–276. doi:10.1038/ng.1073.

Khansefid, M., Pryce, J. E., Bolormaa, S., Miller, S. P., Wang, Z., Li, C., et al. (2014). Estimation of genomic breeding values for residual feed intake in a multibreed cattle population. *J. Anim. Sci*. 92, 3270–3283. doi:10.2527/JAS.2014-7375.

Killinger, K. M., Calkins, C. R., Umberger, W. J., Feuz, D. M., and Eskridge, K. M. (2004). Consumer sensory acceptance and value for beef steaks of similar tenderness, but differing in marbling level. *J. Anim. Sci*. 82, 3294–3301. doi:10.2527/2004.82113294X.

Kim, M., Oh, I., and Ahn, J. (2018). An Improved Method for Prediction of Cancer Prognosis by Network Learning. *Genes (Basel)*. 9, 478. doi:10.3390/GENES9100478.

Kirchhoff, J., Uhlenbruck, S., Goris, K., Keil, G. M., and Herrler, G. (2014). Three viruses of the bovine respiratory disease complex apply different strategies to initiate infection. *Vet. Res*. 45, 20. doi:10.1186/1297-9716-45-20.

Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., et al. (2014). Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *J. Clin. Microbiol*. 52, 438–448. doi:10.1128/JCM.02485-13.

Koch, R. M., Swiger, L. A., Chambers, D., and Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *J. Anim. Sci*. 22, 486–494. doi:10.2527/jas1963.222486x.

Kreimer, A., and Pe'er, I. (2013). Variants in exons and in transcription factors affect gene expression in trans. *Genome Biol*. 14, 1–13. doi:10.1186/GB-2013-14-7-R71.

Kronenberg, F. (2012). "Metabolic Traits as Intermediate Phenotypes," in Genetics Meets Metabolomics: From Experiment to Systems Biology (Springer, New York, NY), 255–264. doi:10.1007/978-1-4614-1689-0_15.

Lacal, J. C. (2001). Choline kinase: a novel target for antitumor drugs. *IDrugs* 4, 419–26. doi:10.1042/bj1800559.

Lande, R., and Thompson, R. (1990). Efficiency of Marker-Assisted Selection in the Improvement of Quantitative Traits. *Genetics* 124, 756. doi:10.1093/genetics/124.3.743.

Larson, R. L., and Step, D. L. (2012). Evidence-Based Effectiveness of Vaccination Against Mannheimia haemolytica, Pasteurella multocida, and Histophilus somni in Feedlot Cattle for Mitigating the Incidence and Effect of Bovine Respiratory Disease Complex. *Vet. Clin. North Am. Food Anim. Pract*. 28, 97–106. doi:10.1016/j.cvfa.2011.12.005.

Lee, H. J., Park, H. S., Kim, W., Yoon, D., and Seo, S. (2014). Comparison of metabolic network between muscle and intramuscular adipose tissues in Hanwoo beef cattle using a systems biology approach. *Int. J. Genomics* 2014, 679437. doi:10.1155/2014/679437.

Lee, M. Y., Kim, T.-K., Walters, K.-A., and Wang, K. (2019). A biological function based biomarker panel optimization process. *Sci. Rep*. 9, 7365. doi:10.1038/s41598-019-43779-2.

Li, B., Zhang, N., Wang, Y.-G., George, A. W., Reverter, A., and Li, Y. (2018). Genomic Prediction of Breeding Values Using a Subset of SNPs Identified by Three Machine Learning Methods. *Front. Genet*. 0, 237. doi:10.3389/FGENE.2018.00237.

Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., et al. (2019). Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. *Microbiome* 7, 1–17. doi:10.1186/s40168-019-0699-1.

Li, J., Akanno, E. C., Valente, T. S., Abo-Ismail, M., Karisa, B. K., Wang, Z., et al. (2020). Genomic heritability and genome-wide association studies of plasma metabolites in crossbred beef cattle. *Front. Genet*. 11, 538600. doi:10.3389/fgene.2020.538600.

Li, J., Mukiibi, R., Wang, Y., Plastow, G. S., and Li, C. (2021). Identification of candidate genes and enriched biological functions for feed efficiency traits by integrating plasma metabolites and imputed whole genome sequence variants in beef cattle. *BMC Genomics* 22, 823. doi:10.1186/S12864-021-08064-5.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/BIOINFORMATICS/BTT656.

Lin, Y., Zhang, W., Cao, H., Li, G., and Du, W. (2020). Classifying Breast Cancer Subtypes Using Deep Neural Networks Based on Multi-Omics Data. *Genes (Basel)*. 11, 888. doi:10.3390/GENES11080888.

López-López, Á., López-Gonzálvez, Á., Barker-Tejeda, T. C., and Barbas, C. (2018). A review of validated biomarkers obtained through metabolomics. *Expert Rev. Mol. Diagn*. 18, 557–575. doi:10.1080/14737159.2018.1481391.

Lourenco, D. A. L., Tsuruta, S., Fragomeni, B. O., Masuda, Y., Aguilar, I., Legarra, A., et al. (2015). Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J. Anim. Sci*. 93, 2653–2662. doi:10.2527/jas.2014-8836.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15, 550. doi:10.1186/S13059-014-0550-8.

Lu, D., Akanno, E. C., Crowley, J. J., Schenkel, F., Li, H., De Pauw, M., et al. (2016). Accuracy of genomic predictions for feed efficiency traits of beef cattle using 50K and imputed HD genotypes. *J. Anim. Sci*. 94, 1342–1353. doi:10.2527/JAS.2015-0126.

Lu, D., Miller, S., Sargolzaei, M., Kelly, M., Vander Voort, G., Caldwell, T., et al. (2013). Genome-wide association analyses for growth and feed efficiency traits in beef cattle. *J. Anim. Sci*. 91, 3612–3633. doi:10.2527/jas.2012-5716.

Lu, D., Sargolzaei, M., Kelly, M., Li, C., Voort, G. Vander, Wang, Z., et al. (2012). Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Genet.* 3, 152. doi:10.3389/fgene.2012.00152.

Lu, D., Sargolzaei, M., Kelly, M., Vander Voort, G., Wang, Z., Mandell, I., et al. (2013). Genome-wide association analyses for carcass quality in crossbred beef cattle. *BMC Genet*. 14, 80. doi:10.1186/1471-2156-14-80.

Ma, H., Lu, F., Guo, Y., Shen, Z., and Chen, Y. (2021). The Prognostic Value of Leucine-Rich α 2 Glycoprotein 1 in Pediatric Spinal Cord Injury. *Biomed Res. Int*. 2021, 7365204. doi:10.1155/2021/7365204.

Ma, S., Ren, J., and Fenyö, D. (2016). Breast Cancer Prognostics Using Multi-Omics Data. *AMIA Jt. Summits Transl. Sci. Proc*. 2016, 52–59.

MacLeod, I. M., Bowman, P. J., Vander Jagt, C. J., Haile-Mariam, M., Kemper, K. E., Chamberlain, A. J., et al. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* 17, 144. doi:10.1186/S12864-016-2443-6.

Manoli, I., and Venditti, C. P. (2016). Disorders of branched chain amino acid metabolism. *Transl. Sci. Rare Dis*. 1, 91–110. doi:10.3233/trd-160009.

Mantovani, A., Cassatella, M. A., Costantini, C., and Jaillon, S. (2011). Neutrophils in the activation and regulation of innate and adaptive immunity. *Nat. Rev. Immunol*. 11, 519–531. doi:10.1038/nri3024.

Mao, F., Chen, L., Vinsky, M., Okine, E., Wang, Z., Basarab, J., et al. (2013). Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci*. 91, 2067–2076. doi:10.2527/jas.2012-5470.

Marengo-Rowe, A. J. (2006). Structure-function relations of human hemoglobins. Proc. (Bayl. Univ. Med. Cent). 19, 239–245. doi:10.1080/08998280.2006.11928171.

Marete, A., Sahana, G., Fritz, S., Lefebvre, R., Barbat, A., Lund, M. S., et al. (2018). Genome-wide association study for milking speed in French Holstein cows. *J. Dairy Sci*. 101, 6205–6219. doi:10.3168/jds.2017-14067.

Matthews, J. O., Southern, L. L., Higbie, A. D., Persica, M. A., and Bidner, T. D. (2001). Effects of betaine on growth, carcass characteristics, pork quality, and plasma metabolites of finishing pigs. *J. Anim. Sci*. 79, 722–728. doi:10.2527/2001.793722x.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 40, 4288. doi:10.1093/NAR/GKS042.

McClellan, J., and King, M.-C. (2010). Genetic Heterogeneity in Human Disease. *Cell* 141, 210–217. doi:10.1016/J.CELL.2010.03.032.

McKenna, C., Keogh, K., Porter, R. K., Waters, S. M., Cormican, P., and Kenny, D. A. (2021). An examination of skeletal muscle and hepatic tissue transcriptomes from beef cattle divergent for residual feed intake. *Sci. Rep*. 11, 8942. doi:10.1038/s41598-021-87842-3.

McKenna, C., Porter, R. K., Keogh, K. A., Waters, S. M., McGee, M., and Kenny, D. A. (2018). Residual feed intake phenotype and gender affect the expression of key genes of the lipogenesis pathway in subcutaneous adipose tissue of beef cattle. *J. Anim. Sci. Biotechnol*. 9, 68. doi:10.1186/s40104-018-0282-9.

McKeown, L., Aalhus, J., Larsen, I., Stothard, P., and Wang, Z. (2013). Bridging the "Phenomic Gap": Creation of a database containing phenotypes and genotypes for economically important traits for beef cattle. Edmonton, Canada. Final Report to the Alberta Livestock and Meat Agency.

Medeiros de Oliveira Silva, R., Bonvino Stafuzza, N., de Oliveira Fragomeni, B., Miguel Ferreira de Camargo, G., Matos Ceacero, T., Noely dos Santos Gonçalves Cyrillo, J., et al. (2017). Genome-wide association study for carcass traits in an experimental Nelore cattle population. *PLoS One* 12, e0169860. doi:10.1371/journal.pone.0169860.

Medzhitov, R. (2010). Inflammation 2010: New Adventures of an Old Flame. *Cell* 140, 771–776. doi:10.1016/j.cell.2010.03.006.

Mehrban, H., Lee, D. H., Moradi, M. H., Ilcho, C., Naserkheil, M., and Ibáñez-Escriche, N. (2017). Predictive performance of genomic selection methods for carcass traits in Hanwoo beef cattle: Impacts of the genetic architecture. *Genet. Sel. Evol*. 49, 1. doi:10.1186/s12711-016-0283-0.

Melzer, N., Wittenburg, D., and Repsilber, D. (2013). Integrating Milk Metabolite Profile Information for the Prediction of Traditional Milk Traits Based on SNP Information for Holstein Cows. *PLoS One* 8, e70256. doi:10.1371/journal.pone.0070256.

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi:10.1093/GENETICS/157.4.1819.

Meuwissen, T., Hayes, B., and Goddard, M. (2013). Accelerating improvement of livestock with genomic selection. *Annu. Rev. Anim. Biosci*. 1, 221–237. doi:10.1146/annurev-animal-031412-103705.

Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Anim. Front*. 6, 6–14. doi:10.2527/AF.2016-0002.

Meuwissen, T., van den Berg, I., and Goddard, M. (2021). On the use of whole-genome sequence data for across-breed genomic prediction and fine-scale mapping of QTL. *Genet. Sel. Evol*. 53, 19. doi:10.1186/s12711-021-00607-4.

Mierziak, J., Burgberger, M., and Wojtasik, W. (2021). 3-hydroxybutyrate as a metabolite and a signal molecule regulating processes of living organisms. *Biomolecules* 11, 402. doi:10.3390/biom11030402.

Milatovich, A., Kitamura, T., Miyajima, A., and Francke, U. (1993). Gene for the alpha-subunit of the human interleukin-3 receptor (IL3RA) localized to the X-Y pseudoautosomal region. *Am. J. Hum. Genet*. 53, 1146.

Min, S., Lee, B., and Yoon, S. (2017). Deep learning in bioinformatics. *Brief. Bioinform*. 18, 851–869. doi:10.1093/BIB/BBW068.

Mirza, B., Wang, W., Wang, J., Choi, H., Chung, N. C., and Ping, P. (2019). Machine Learning and Integrative Analysis of Biomedical Big Data. *Genes (Basel)*. 10, 87. doi:10.3390/GENES10020087.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., and Lee, D. H. (2002). BLUPF90 and related programs (BGF90). in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, 21–22.

Moisá, S. J., Aly, S. S., Lehenbauer, T. W., Love, W. J., Rossitto, P. V., Eenennaam, A. L. Van, et al. (2019). Association of plasma haptoglobin concentration and other biomarkers with bovine respiratory disease status in pre-weaned dairy calves. *J. Vet. Diagn. Invest*. 31, 40. doi:10.1177/1040638718807242.

Montgomery, S. P., Sindt, J. J., Greenquist, M. A., Miller, W. F., Pike, J. N., Loe, E. R., et al. (2009). Plasma metabolites of receiving heifers and the relationship between apparent bovine respiratory disease, body weight gain, and carcass characteristics. *J. Anim. Sci*. 87, 328–333. doi:10.2527/jas.2008-0969.

Morgavi, D. P., Kelly, W. J., Janssen, P. H., and Attwood, G. T. (2013). "Rumen microbial (meta)genomics and its application to ruminant production," in Animal (Cambridge University Press), 184–201. doi:10.1017/S1751731112000419.

Muggli-Cockett, N. E., Cundiff, L. V., and Gregory, K. E. (1992). Genetic analysis of bovine respiratory disease in beef calves during the first year of life1. *J. Anim. Sci*. 70, 2013–2019. doi:10.2527/1992.7072013x.

Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Grant, J. R., Mah, J., Wang, Z., et al. (2011). Associations of marker panel scores with feed intake and efficiency traits in beef cattle using preselected single nucleotide polymorphisms. *J. Anim. Sci*. 89, 3362–3371. doi:10.2527/JAS.2010-3362.

Mukiibi, R. (2019). Evaluation of genomic prediction accuracy of feed efficiency and the related traits with integration of differentially expressed gene SNPs.

Mukiibi, R., Vinsky, M., Keogh, K. A., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2018). Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. *Sci. Rep*. 8, 7303. doi:10.1038/s41598-018-25605-3.

Mukiibi, R., Vinsky, M., Keogh, K. A., Stothard, P., Waters, S. M., and Li, C. (2018). Transcriptome analysis of liver tissues in Charolais steers with divergent marbling phenotypes. in Proceedings of the World Congress on Genetics Applied to Livestock Production (Auckland, New Zealand), 11.626.

Mukiibi, R., Vinsky, M., Keogh, K., Fitzsimmons, C., Stothard, P., Waters, S. M., et al. (2019). Liver transcriptome profiling of beef steers with divergent growth rate, feed intake, or metabolic body weight phenotypes. *J. Anim. Sci*. 97, 4386–4404. doi:10.1093/jas/skz315.

N'jai, A. U., Rivera, J., Atapattu, D. N., Owusu-Ofori, K., and Czuprynski, C. J. (2013). Gene expression profiling of bovine bronchial epithelial cells exposed in vitro to bovine herpesvirus 1 and Mannheimia haemolytica. *Vet. Immunol. Immunopathol*. 155, 182–189. doi:10.1016/j.vetimm.2013.06.012.

Nalpas, N. C., Park, S. D. E., Magee, D. A., Taraktsoglou, M., Browne, J. A., Conlon, K. M., et al. (2013). Whole-transcriptome, high-throughput RNA sequence analysis of the bovine macrophage response to Mycobacterium bovis infection in vitro. *BMC Genomics* 14, 230. doi:10.1186/1471-2164-14-230/TABLES/2.

Neibergs, H. L., Neibergs, J. S., Wojtowicz, A. J., Taylor, J. F., Seabury, C. M., and Womack, J. E. (2014a). Economic benefits of using genetic sele ction to reduce the prev alence of bovine respiratory disease complex in beef feedlot cattle. in Proceedings of the Beef Improvement Federation (BIF) Annual Meeting and Research Symposium, 82–86.

Neibergs, H. L., Seabury, C. M., Wojtowicz, A. J., Wang, Z., Scraggs, E., Kiser, J. N., et al. (2014). Susceptibility loci revealed for bovine respiratory disease complex in pre-weaned holstein calves. *BMC Genomics* 15, 1164. doi:10.1186/1471-2164-15-1164.

Nicholson, G., Rantalainen, M., Li, J. V., Maher, A. D., Malmodin, D., Ahmadi, K. R., et al. (2011a). A genome-wide metabolic QTL analysis in europeans implicates two Loci shaped by recent positive selection. *PLoS Genet*. 7, e1002270. doi:10.1371/journal.pgen.1002270.

Nicholson, G., Rantalainen, M., Maher, A. D., Li, J. V., Malmodin, D., Ahmadi, K. R., et al. (2011b). Human metabolic profiles are stably controlled by genetic and environmental variation. *Mol. Syst. Biol*. 7, 525. doi:10.1038/msb.2011.57.

Nicholson, J. K., and Lindon, J. C. (2008). Systems biology: Metabonomics. *Nature* 455, 1054–1056. doi:10.1038/4551054a.

Nielsen, M. K., MacNeil, M. D., Dekkers, J. C. M., Crews, D. H., Rathje, T. A., Enns, R. M., et al. (2013). Life-cycle, total-industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation11The development of this commentary was supported by the Beef Improvement Federation. *Prof. Anim. Sci*. 29, 559–565. doi:10.15232/S1080-7446(15)30285-0.

Nishino, N., and Shinde, S. (2007). Ethanol and 2,3-butanediol production in whole-crop rice silage. *Grassl. Sci*. 53, 196–198. doi:10.1111/j.1744-697x.2007.00089.x.

Nkrumah, J. D., Basarab, J. A., Wang, Z., Li, C., Price, M. A., Okine, E. K., et al. (2007). Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. *J. Anim. Sci*. 85, 2711–2720. doi:10.2527/jas.2006-767.

Nkrumah, J. D., Keisler, D. H., Crews, D. H., Basarab, J. A., Wang, Z., Li, C., et al. (2007a). Genetic and phenotypic relationships of serum leptin concentration with performance, efficiency of gain, and carcass merit of feedlot cattle1. *J. Anim. Sci*. 85, 2147–2155. doi:10.2527/jas.2006-764.

Nkrumah, J. D., Okine, E. K., Mathison, G. W., Schmid, K., Li, C., Basarab, J. A., et al. (2006). Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J. Anim. Sci*. 84, 145–153. doi:10.2527/2006.841145x.

Nkrumah, J. D., Sherman, E. L., Li, C., Marques, E., Crews, D. H., Bartusiak, R., et al. (2007b). Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *J. Anim. Sci*. 85, 3170–3181. doi:10.2527/jas.2007-0234.

Novais, F. J., Pires, P. R. L., Alexandre, P. A., Dromms, R. A., Iglesias, A. H., Ferraz, J. B. S., et al. (2019). Identification of a metabolomic signature associated with feed efficiency in beef cattle. *BMC Genomics* 20, 8. doi:10.1186/s12864-018-5406-2.

Obitsu, T., Nishimura, K., Udaka, Y., Sugino, T., and Taniguchi, K. (2013). "Effects of ethanol on splanchnic nutrient metabolism in sheep at different intake levels," in Energy and protein metabolism and nutrition in sustainable animal production (Wageningen Academic Publishers), 441–442. doi:10.3920/978-90-8686-781-3_164.

Ogawa, S., Matsuda, H., Taniguchi, Y., Watanabe, T., Kitamura, Y., Tabuchi, I., et al. (2017). Genomic prediction for carcass traits in Japanese Black cattle using single nucleotide polymorphism markers of different densities. *Anim. Prod. Sci*. 57, 1631. doi:10.1071/AN15696.

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., et al. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 49, W388–W396. doi:10.1093/NAR/GKAB382.

Patel, S. S., Molnar, M. Z., Tayek, J. A., Ix, J. H., Noori, N., Benner, D., et al. (2013). Serum creatinine as a marker of muscle mass in chronic kidney disease: Results of a cross-sectional study and review of literature. *J. Cachexia. Sarcopenia Muscle* 4, 19–29. doi:10.1007/s13539-012-0079-1.

Picard, M., Scott-Boyer, M.-P., Bodein, A., Périn, O., and Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. *Comput. Struct. Biotechnol. J*. 19, 3735–3746. doi:10.1016/J.CSBJ.2021.06.030.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Auwera, G. A. Van der, et al. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178. doi:10.1101/201178.

Preston, G. A., and Weinberger, D. R. (2005). Intermediate phenotypes in schizophrenia: a selective review. *Dialogues Clin. Neurosci*. 7, 165. doi:10.31887/DCNS.2005.7.2/GPRESTON.

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

Ragvin, A., Moro, E., Fredman, D., Navratilova, P., Drivenes, Ø., Engström, P. G., et al. (2010). Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc. Natl. Acad. Sci*. 107, 775–780. doi:10.1073/PNAS.0911591107.

Ramsey, R., Doye, D., Ward, C., McGrann, J., Falconer, L., and Bevers, S. (2005). Factors affecting beef cow-herd costs, production, and profits. *J. Agric. Appl. Econ*. 37, 91–99. doi:10.1017/s1074070800007124.

Rice, J. A., Carrasco-Medina, L., Hodgins, D. C., and Shewen, P. E. (2008). Mannheimia haemolytica and bovine respiratory disease. *Anim. Heal. Res. Rev*. 8, 117–128. doi:10.1017/S1466252307001375.

Richardson, E. C., and Herd, R. M. (2004). Biological basis for variation in residual feed intake in beef cattle. 2. Synthesis of results following divergent selection. *Aust. J. Exp. Agric*. 44, 431–440. doi:10.1071/ea02221.

Risso, D., Ngai, J., Speed, T. P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol*. 32, 896–902. doi:10.1038/nbt.2931.

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 11, R25. doi:10.1186/GB-2010-11-3-R25.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.

Ron-Harel, N., Ghergurovich, J. M., Notarangelo, G., LaFleur, M. W., Tsubosaka, Y., Sharpe, A. H., et al. (2019). T Cell Activation Depends on Extracellular Alanine. *Cell Rep*. 28, 3011-3021.e4. doi:10.1016/j.celrep.2019.08.034.

Rowe, H. C., Hansen, B. G., Halkier, B. A., and Kliebenstein, D. J. (2008). Biochemical networks and epistasis shape the Arabidopsis thaliana metabolome. *Plant Cell* 20, 1199–1216. doi:10.1105/tpc.108.058131.

Sarup, P., Jensen, J., Ostersen, T., Henryon, M., and Sørensen, P. (2016). Increased prediction accuracy using a genomic feature model including prior information on quantitative trait locus regions in purebred Danish Duroc pigs. *BMC Genet*. 17, 11. doi:10.1186/s12863-015-0322-9.

Schiopu, A., and Cotoi, O. S. (2013). S100A8 and S100A9: DAMPs at the crossroads between innate immunity, traditional risk factors, and cardiovascular disease. *Mediators Inflamm*. 2013, 828354. doi:10.1155/2013/828354.

Schlosser, P., Li, Y., Sekula, P., Raffler, J., Grundner-Culemann, F., Pietzner, M., et al. (2020). Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. *Nat. Genet*. 52, 167–176. doi:10.1038/s41588-019-0567-8.

Schneider, M. J., Tait, R. G., Ruble, M. V., Busby, W. D., and Reecy, J. M. (2010). Evaluation of fixed sources of variation and estimation of genetic parameters for incidence of bovine respiratory disease in preweaned calves and feedlot cattle. *J. Anim. Sci*. 88, 1220–1228. doi:10.2527/jas.2008-1755.

Schröder, U. J., and Staufenbiel, R. (2006). Invited Review: Methods to Determine Body Fat Reserves in the Dairy Cow with Special Regard to Ultrasonographic Measurement of Backfat Thickness. *J. Dairy Sci*. 89, 1–14. doi:10.3168/JDS.S0022-0302(06)72064-1.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole blood transcriptomic analysis of beef cattle at arrival identifies potential

predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. *PLoS One* 15, e0227507. doi:10.1371/JOURNAL.PONE.0227507.

Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2021). Comprehensive at-arrival transcriptomic analysis of post-weaned beef cattle uncovers type I interferon and antiviral mechanisms associated with bovine respiratory disease mortality. *PLoS One* 16, e0250758. doi:10.1371/JOURNAL.PONE.0250758.

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., et al. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18, 386. doi:10.1186/s12864-017-3754-y.

Selevsek, N., Chang, C. Y., Gillet, L. C., Navarro, P., Bernhardt, O. M., Reiter, L., et al. (2015). Reproducible and Consistent Quantification of the Saccharomyces cerevisiae Proteome by SWATH-mass spectrometry. *Mol. Cell. Proteomics* 14, 739–749. doi:10.1074/MCP.M113.035550.

Shabalin, A. A. (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 28, 1353–1358. doi:10.1093/BIOINFORMATICS/BTS163.

Sharifi-Noghabi, H., Zolotareva, O., Collins, C. C., and Ester, M. (2019). MOLI: multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* 35, i501–i509. doi:10.1093/BIOINFORMATICS/BTZ318.

Shin, S.-Y., Fauman, E. B., Petersen, A.-K., Krumsiek, J., Santos, R., Huang, J., et al. (2014). An atlas of genetic influences on human blood metabolites. *Nat. Genet*. 46, 543–550. doi:10.1038/ng.2982.

Silva, R. M. O., Fragomeni, B. O., Lourenco, D. A. L., Magalhães, A. F. B., Irano, N., Carvalheiro, R., et al. (2016). Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. *J. Anim. Sci*. 94, 3613–3623. doi:10.2527/JAS.2016-0401.

Smith, R. A. (1998). Impact of disease on feedlot performance: a review. *J. Anim. Sci*. 76, 272–274. doi:10.2527/1998.761272X.

Smith, R. A. (2000). Effects of feedlot disease on economics, production and carcass value. in The American Association of Bovine Practitioners Proceedings, 125–128.

Smith, Z. K., and Johnson, B. J. (2020). Mechanisms of steroidal implants to improve beef cattle growth: a review. *J. Appl. Anim. Res*. 48, 133–141. doi:10.1080/09712119.2020.1751642.

Snelling, W. M., Cushman, R. A., Fortes, M. R. S., Reverter, A., Bennett, G. L., Keele, J. W., et al. (2012). PHYSIOLOGY AND ENDOCRINOLOGY SYMPOSIUM: How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. *J. Anim. Sci*. 90, 1152–1165. doi:10.2527/JAS.2011-4581.

Snelling, W. M., Cushman, R. A., Keele, J. W., Maltecca, C., Thomas, M. G., Fortes, M. R. S., et al. (2013). BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection. *J. Anim. Sci*. 91, 537–552. doi:10.2527/JAS.2012-5784.

Snowder, G. (2009). Genetics, environment and bovine respiratory disease. *Anim. Health Res. Rev*. 10, 117–119. doi:10.1017/S1466252309990144.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2005). Influence of breed, heterozygosity, and disease incidence on estimates of variance components of respiratory disease in preweaned beef calves. *J. Anim. Sci*. 83, 1247–1261. doi:10.2527/2005.8361247x.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2006). Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *J. Anim. Sci*. 84, 1999–2008. doi:10.2527/jas.2006-046.

Song, M., Greenbaum, J., Luttrell, J., IV, Zhou, W., Wu, C., et al. (2020). A Review of Integrative Imputation for Multi-Omics Datasets. *Front. Genet*. 11, 570255. doi:10.3389/FGENE.2020.570255.

Srikanth, K., Lee, S.-H., Chung, K.-Y., Park, J.-E., Jang, G.-W., Park, M.-R., et al. (2020). A Gene-Set Enrichment and Protein–Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. *Genes (Basel)*. 11, 316. doi:10.3390/GENES11030316.

Stanford, K., Zaheer, R., Klima, C., McAllister, T., Peters, D., Niu, Y. D., et al. (2020). Antimicrobial resistance in members of the bacterial bovine respiratory disease complex isolated from lung tissue of cattle mortalities managed with or without the use of antimicrobials. *Microorganisms* 8, 288. doi:10.3390/microorganisms8020288.

Stranger, B. E., Stahl, E. A., and Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 187, 367–383. doi:10.1534/genetics.110.120907.

Suhre, K., and Gieger, C. (2012). Genetic variation in metabolic phenotypes: study designs and applications. *Nat. Rev. Genet*. 13, 759–769. doi:10.1038/nrg3314.

Suhre, K., Shin, S. Y., Petersen, A. K., Mohney, R. P., Meredith, D., Wägele, B., et al. (2011). Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 477, 54–62. doi:10.1038/nature10354.

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., et al. (2018). Genomic atlas of the human plasma proteome. *Nature* 558, 73–79. doi:10.1038/s41586-018-0175-2.

Sun, H. Z., Srithayakumar, V., Jiminez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics* 112, 3968–3977. doi:10.1016/J.YGENO.2020.07.014.

Sun, Y. V., and Hu, Y. J. (2016). Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv. Genet*. 93, 147–190. doi:10.1016/BS.ADGEN.2015.11.004.

Tanha, H. M., Sathyanarayanan, A., and Nyholt, D. R. (2021). Genetic overlap and causality between blood metabolites and migraine. *Am. J. Hum. Genet*. 108, 2086–2098. doi:10.1016/J.AJHG.2021.09.011.

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., and Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *Can. Vet. J*. 51, 1095–1102.

Theurer, M. E., Larson, R. L., and White, B. J. (2015). Systematic review and meta-analysis of the effectiveness of commercially available vaccines against bovine herpesvirus, Bovine viral diarrhea virus, Bovine respiratory syncytial virus, And parainfluenza type 3 virus for mitigation of bovine respiratory di. *J. Am. Vet. Med. Assoc*. 246, 126–142. doi:10.2460/javma.246.1.126.

Tizioto, P. C., Coutinho, L. L., Decker, J. E., Schnabel, R. D., Rosa, K. O., Oliveira, P. S. N., et al. (2015). Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. *BMC Genomics* 16, 242. doi:10.1186/s12864-015-1464-x.

Tizioto, P. C., Kim, J. W., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015). Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: An RNA-sequence analysis of the bronchial lymph node transcriptome. *PLoS One* 10, e0131459. doi:10.1371/journal.pone.0131459.

Todd, D. L., Roughsedge, T., and Woolliams, J. A. (2014). Genomic selection using beef commercial carcass phenotypes. *Animal* 8, 388–394. doi:10.1017/S1751731113002231.

Torres-Vázquez, J. A., van der Werf, J. H. J., and Clark, S. A. (2018). Genetic and phenotypic associations of feed efficiency with growth and carcass traits in Australian Angus cattle. *J. Anim. Sci*. 96, 4521–4531. doi:10.1093/jas/sky325.

Turner, S. D. (2014). qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*, 005165. doi:10.1101/005165.

USDA (2011). Part IV: Health and Health Management on US Feedlots with a Capacity of 1,000 or More Head. Available at: https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_PartIV.pdf.

Van der Auwera, G. A., and O'Connor, B. D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. 1st Editio. O'Reilly Media, Inc.

van der Veen, J. N., Kennelly, J. P., Wan, S., Vance, J. E., Vance, D. E., and Jacobs, R. L. (2017). The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim. Biophys. Acta - Biomembr*. 1859, 1558–1572. doi:10.1016/j.bbamem.2017.04.006.

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

Wang, H., Misztal, I., Aguilar, I., Legarra, A., and Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res. (Camb)*. 94, 73–83. doi:10.1017/S0016672312000274.

Wang, X., and Kadarmideen, H. N. (2019). Metabolomics analyses in high-low feed efficient dairy cows reveal novel biochemical mechanisms and predictive biomarkers. *Metabolites* 9, 151. doi:10.3390/metabo9070151.

Wang, X., Wang, K., Radovich, M., Wang, Y., Wang, G., Feng, W., et al. (2009). Genome-wide prediction of cis-acting RNA elements regulating tissue-specific pre-mRNA alternative splicing. *BMC Genomics* 10, S4. doi:10.1186/1471-2164-10-S1-S4.

Wang, X., Zhang, Y., Zhang, X., Wang, D., Jin, G., Li, B., et al. (2017). The comprehensive liver transcriptome of two cattle breeds with different intramuscular fat content. *Biochem. Biophys. Res. Commun*. 490, 1018–1025. doi:10.1016/j.bbrc.2017.06.157.

Wang, Y., Zhang, F., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (2020). Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. *BMC Genomics* 21, 38. doi:10.1186/s12864-019-6273-1.

Weber, K. L., Welly, B. T., Van Eenennaam, A. L., Young, A. E., Porto-Neto, L. R., Reverter, A., et al. (2016). Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. *PLoS One* 11, e0152274. doi:10.1371/journal.pone.0152274.

Wen, W., Li, D., Li, X., Gao, Y., Li, W., Li, H., et al. (2014). Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun*. 5, 3438. doi:10.1038/ncomms4438.

Widmann, P., Reverter, A., Fortes, M. R. S., Weikard, R., Suhre, K., Hammon, H., et al. (2013). A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* 14, 798. doi:10.1186/1471-2164-14-798.

Wimmer, V., Albrecht, T., Auinger, H. J., and Schön, C. C. (2012). Synbreed: A framework for the analysis of genomic prediction data using R. *Bioinformatics* 28, 2086–2087. doi:10.1093/bioinformatics/bts335.

Wu, J., Yin, H., Zhu, J., Buckanovich, R. J., Thorpe, J. D., Dai, J., et al. (2015). Validation of LRG1 as a Potential Biomarker for Detection of Epithelial Ovarian Cancer by a Blinded Study. *PLoS One* 10, e0121112. doi:10.1371/JOURNAL.PONE.0121112.

Wu, P., Wang, K., Zhou, J., Chen, D., Jiang, A., Jiang, Y., et al. (2021). A combined GWAS approach reveals key loci for socially-affected traits in Yorkshire pigs. *Commun. Biol*. 4, 891. doi:10.1038/s42003-021-02416-3.

Wyss, M., and Kaddurah-Daouk, R. (2000). Creatine and creatinine metabolism. *Physiol. Rev*. 80, 1107–1213. doi:10.1152/physrev.2000.80.3.1107.

Xia, Y., Chen, S., Zhu, G., Huang, R., Yin, Y., and Ren, W. (2018). Betaine inhibits interleukin-1β production and release: Potential mechanisms. *Front. Immunol*. 9, 2670. doi:10.3389/fimmu.2018.02670.

Xu, L., Gao, N., Wang, Z., Xu, L., Liu, Y., Chen, Y., et al. (2020). Incorporating genome annotation into genomic prediction for carcass traits in Chinese Simmental beef cattle. *Front. Genet*. 11, 481. doi:10.3389/fgene.2020.00481.

Xu, P., Liu, P., Zhou, C., Shi, Y., Wu, Q., Yang, Y., et al. (2019). A Multi-Omics Study of Chicken Infected by Nephropathogenic Infectious Bronchitis Virus. *Viruses* 11, 1070. doi:10.3390/V11111070.

Yan, J., Risacher, S. L., Shen, L., and Saykin, A. J. (2018). Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. *Brief. Bioinform*. 19, 1370–1381. doi:10.1093/bib/bbx066.

Yang, F. J., Hsieh, C. Y., Shu, K. H., Chen, I. Y., Pan, S. Y., Chuang, Y. F., et al. (2020). Plasma Leucine-Rich α-2-Glycoprotein 1 Predicts Cardiovascular Disease Risk in End-Stage Renal Disease. *Sci. Rep*. 10, 5988. doi:10.1038/s41598-020-62989-7.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet*. 88, 76–82. doi:10.1016/j.ajhg.2010.11.011.

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet*. 46, 100–106. doi:10.1038/ng.2876.

Ye, S., Li, J., and Zhang, Z. (2020). Multi-omics-data-assisted genomic feature markers preselection improves the accuracy of genomic prediction. *J. Anim. Sci. Biotechnol*. 11, 109. doi:10.1186/S40104-020-00515-5.

Yousri, N. A., Kastenmüller, G., Gieger, C., Shin, S.-Y., Erte, I., Menni, C., et al. (2014). Long term conservation of human metabolic phenotypes and link to heritability. *Metabolomics* 10, 1005–1017. doi:10.1007/s11306-014-0629-y.

Yu, Y., and Song, G. (2020). "Lipopolysaccharide-binding protein and bactericidal/permeability-increasing protein in lipid metabolism and cardiovascular diseases," in Lipid Transfer in Lipoprotein Metabolism and Cardiovascular Disease, ed. X. Jiang (Springer, Singapore), 27–35. doi:10.1007/978-981-15-6082-8_3.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res*. 46, D754–D761. doi:10.1093/nar/gkx1098.

Zhang, F., Wang, Y., Mukiibi, R., Chen, L., Vinsky, M., Plastow, G., et al. (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, 36. doi:10.1186/s12864-019-6362-1.

Zhang, H., Yin, L., Wang, M., Yuan, X., and Liu, X. (2019). Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Front. Genet*. 10, 189. doi:10.3389/fgene.2019.00189.

Zhang, Y., Parmigiani, G., and Johnson, W. E. (2020b). ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma*. 2, lqaa078. doi:10.1093/NARGAB/LQAA078.

Zhang, Y., Wang, Y., Li, Y., Wu, J., Wang, X., Bian, C., et al. (2020c). Genome-wide association study reveals the genetic determinism of growth traits in a Gushi-Anka F2 chicken population. *Heredity (Edinb)*. 126, 293–307. doi:10.1038/s41437-020-00365-x.

Zhao, S., and Li, L. (2020). Chemical derivatization in LC-MS-based metabolomics study. *TrAC Trends Anal. Chem*. 131, 115988. doi:10.1016/J.TRAC.2020.115988.

Zhao, S., Jiang, H., Liang, Z.-H., and Ju, H. (2020). Integrating Multi-Omics Data to Identify Novel Disease Genes and Single-Neucleotide Polymorphisms. *Front. Genet*. 0, 1336. doi:10.3389/FGENE.2019.01336.

Zhao, W., Rasheed, A., Tikkanen, E., Lee, J.-J., Butterworth, A. S., Howson, J. M., et al. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat. Genet*. 49, 1450–1457. doi:10.1038/NG.3943.

Zhou, G., Li, S., and Xia, J. (2020). "Network-Based Approaches for Multi-omics Integration," in Computational Methods and Data Analysis for Metabolomics (Humana, New York, NY), 469–487. doi:10.1007/978-1-0716-0239-3_23.