Application of Multi-Omics Approaches to Maximize Beef Production

by

Aidin Foroutan Naddafi

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

Doctor of Philosophy

in

Animal Science

Departments of Agricultural, Food & Nutritional Science

and

Biological Sciences

University of Alberta

ABSTRACT

Approximately 70% of the cost of beef production is impacted by dietary intake. Maximizing production efficiency of beef cattle requires not only genetic selection to maximize feed efficiency (i.e. residual feed intake - RFI), but also adequate nutrition throughout all stages of growth and development to maximize efficiency of growth and reproductive capacity - even during gestation. Nutrient restriction during gestation has been shown to negatively affect postnatal growth and development as well as fertility of the offspring. This, when combined with RFI, may significantly affect energy partitioning in the offspring and subsequently important performance traits. Therefore, we decided to conduct a comprehensive multi-omics study (metabolomics, transcriptomics, epigenomics) to understand the biological mechanisms impacted by prenatal nutrition (normal-diet or Ndiet versus low-diet or Ldiet) and/or parental RFI (high-RFI or HRFI versus low-RFI or LRFI) in young Angus bulls. Four different tissues (*Longissimus thoracis* (LT) muscle, semimembranosus (SM) muscle, liver, and testis) and three biofluids (serum, semen, and ruminal fluid) were analyzed. Through the metabolomics study, we created the Bovine Metabolome Database (BMDB; www.bovinedb.ca) which contains 51,801 metabolites with unique compound structures in various tissues and biofluids. We also identified two serum candidate biomarker panels ((1) formate and leucine; (2) C4 (butyrylcarnitine) and LysoPC(28:0)), which can distinguish HRFI from LRFI animals with high sensitivity and specificity (area under the curve from receiver-operator characteristic (ROC) or AUROC > 0.85). Through the transcriptomics study, we found that differences in selection for parental RFI altered gene expression level of myocyte enhancer factor 2A [*MEF2A*] in tissues (LT muscle, SM muscle, liver, and testis) of young Angus bulls. Furthermore, the mRNA abundance of protocadherin 19 [*PCDH19*] in liver, and *MEF2A* in LT muscle were affected by prenatal undernutrition. We also

detected correlations between gene expression in tissues with phenotypic measures of feed efficiency and body weight. Through the epigenomics study, we found 652 and 1400 differentially methylated regions (DMRs) that were affected by maternal diet and parental RFI, respectively. Through pathway analysis of the identified DMRs using the Ingenuity Pathway Analysis (IPA) tool, three networks associated with "cell survival and growth", "disease or abnormalities", and "connective tissue development" were identified as being overrepresented in the DMRs when comparing the Ndiet group to the Ldiet group. Similar pathway analysis for the HRFI and LRFI bulls showed overrepresentation of the number of DMRs in four networks involved in "embryonic development", "DNA replication, DNA repair, and RNA processing", "growth control and homeostasis", as well as "lipid metabolism". These findings provided new knowledge regarding underlying biological mechanisms regulating postnatal responses to prenatal nutrition and feed efficiency in beef bulls.

PREFACE

This thesis is an original work by Aidin Foroutan Naddafi. As detailed in the following, some chapters of this thesis have been submitted for publication as scholarly articles in which Dr. Carolyn Fitzsimmons and Dr. David Scott Wishart were the supervisory authors and have contributed to concepts formation and the manuscript composition.

A version of Chapter 2 was published as

- Aidin Foroutan, Carolyn Fitzsimmons, Rupasri Mandal, Hamed Piri-Moghadam, Jiamin Zheng, AnChi Guo, Carin Li, Le Luo Guan, David S. Wishart. The bovine metabolome. *Metabolites 2020, 10(6), 233; https://doi.org/10.3390/metabo10060233*

A version of Chapter 3 was published as

- Aidin Foroutan, Carolyn Fitzsimmons, Rupasri Mandal, Mark V. Berjanskii, David S. Wishart. Serum metabolite biomarkers for predicting residual feed intake (RFI) of young angus bulls. *Metabolites 2020, 10(12), 491; https://doi.org/10.3390/metabo10120491*

A version of Chapter 4 was published as

- Aidin Foroutan, Julia Devos, David S. Wishart, Changxi Li, Marcos Colazo, John Kastelic, Jacob Thundathil, Carolyn Fitzsimmons. Impact of prenatal nutrition and selection for parental residual feed intake (RFI) on selected gene expression in young Angus bulls. *Livestock Science. 2021, 243, 104365; https://doi.org/10.1016/j.livsci.2020.104365*

CONTRIBUTIONS

Chapter 1: I researched and wrote the entire chapter and generated all the figures used in this chapter. Drs. David Wishart and Carolyn Fitzsimmons provided editing and advice on content and layout.

Chapter 2: I researched and wrote the chapter and generated all of the tables used in this chapter. I performed the literature review, compiled the data and assisted with entering the data into the BMDB. I also assisted in the design, formatting and editing of the BMDB. Dr. Hamed Piri-Moghadam assisted in ICP-MS data acquisition and Jiamin Zheng assisted in LC-MS/MS data acquisition. The other co-authors, Dr. AnChi Guo and Carin Li, performed the programming and database management required for the database while Dr. Carolyn Fitzsimmons, Dr. Rupasri Mandal, Dr. Hamed Piri-Moghadam, Dr. AnChi Guo, Dr. Le Luo Guan, Dr. David Wishart, Jiamin Zheng, and Carin Li provided editing and content advice.

Chapter 3: I researched and wrote the chapter and generated all of the tables used in this chapter. Dr. Mark V. Berjanskii assisted in data analysis. The other co-authors, Dr. Carolyn Fitzsimmons, Dr. Rupasri Mandal, Dr. Mark V. Berjanskii, and Dr. David Wishart provided editing and content advice.

Chapter 4: I researched and wrote the chapter and generated all of the tables used in this chapter. Julia Devos assisted in lab experiments. Dr. Marcos Colazo performed estrous synchronization, artificial insemination and sample collection. Dr. John Kastelic, Dr. Jacob Thundathil, and Dr. Carolyn Fitzsimmons designed the animal experiment and collected tissues from the bull progeny. Dr. Changxi Li calculated all RFI values. Dr. David Wishart, Dr. Changxi Li, Dr. Marcos Colazo, Dr. John Kastelic, Dr. Jacob Thundathil, and Dr. Carolyn Fitzsimmons provided editing and content advice.

Chapter 5: I researched and wrote the chapter and generated all of the tables used in this chapter. Dr. Arun Kommadath converted the subset of filtered Differentially Methylated Region (DMR) coordinates from the bovine UMD3.1.1 assembly to the bovine ARS-UCD1.2 assembly. Dr. Marcos Colazo performed estrous synchronization, artificial insemination and sample collection. Dr. John Kastelic, Dr. Jacob Thundathil, Dr. Carolyn Fitzsimmons, and Chinju Johnson assisted in design and collection of semen for the bull progeny. Dr. Changxi Li calculated all RFI values. Dr. David Wishart and Dr. Carolyn Fitzsimmons provided editing and content advice.

Chapter 6: I researched and wrote the entire chapter. Drs. David Wishart and Carolyn Fitzsimmons provided editing and advice on content and layout.

DEDICATION

This work is dedicated to my wife, Azadeh Yasari, who has been a constant source of support and encouragement in every single moment of my life.

I am truly thankful for having you in my life.

This work is also dedicated to my parents, Dr. Abdolreza Foroutan Naddafi and Mrs. Vida Jahanara, for their love, sacrifice, and tolerance.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude and sincere thanks to my supervisors, Dr. Carolyn Fitzsimmons and Dr. David Wishart, for their patience, professional guidance, and support, during the course of my thesis. Thanks to them, I learned the skills necessary to conduct research in a professional manner. This work would not have been possible without their continuous supervision and help.

I would especially like to thank Dr. Le Luo Guan as a supervisory committee member for her invaluable and constructive comments, suggestions, and criticisms and for the time she spent to guide.

I also would like to thank Dr. Rupasri Mandal, Dr. AnChi Guo, Dr. Hamed Piri-Moghadam, Dr. Mark Berjanskii, Dr. Ghader Manafiazar, Dr. Arun Kommadath, Dr. Allen Zhang, Dr. Changxi Li, Dr. Marcos Colazo, Dr. John Kastelic, Dr. Jacob Thundathil, Julia Devos, Jiamin Zheng, Rahmatollah Rajabzadeh, Amir Behrouzi, Seyed Ali Goldansaz, Shirin Zahraei, Michael Vinsky, Carin Li, Ying (Edison) Dong, Chinju Johnson, Mathew Johnson, and all fellows in The Metabolomics Innovation Centre (TMIC) as well as Alberta Livestock Gentec who have provided me with invaluable assistance, support, and inspiration throughout the course of my study.

I also would like to thank Dr. Burim N. Ametaj, Dr. Fozia Saleem, Dr. Souhaila Bouatra, Dr. Nick Psychogios, Dr. Jianguo Xia, Vanessa Neveu, Roman Eisner, Craig Knox, Umar Farooq, Suzanna M. Dunn, Daya G. V. Emmanuel, Shanthipoosan Sivaraman, Ramprakash P. Periasamy, Michael J. Lewis for building the first platform of the Bovine Metabolome Database.

I highly appreciate the Alberta Livestock Meat Agency (ALMA) for the financial support of this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

1.1. INTRODUCTION

The costs of a cow-calf production cycle are usually divided into six cost categories including: 1) feed and pasture (64%); 2) herd replacement (13%); 3) yardage (13%); 4) labour (6%); 5) marketing (1%); and 6) veterinary (3%) costs [1]. Feed is, by far, the largest financial burden for the beef industry in Western Canada. The cost of feed varies from region-to-region with dependencies on climate and food source availability. It also varies from year-to-year, with a strong dependency on food resource unit costs [1, 2]. In particular, a longer winter-feeding period increases feed costs. To lower winter feed costs and improve profitability, some producers in Western Canada use extended grazing seasons as well as alternative feeds [1]. Other producers provide specific feed/nutritional supplements to various cow/calf groups within the herd (i.e., first calf heifers, mature cows, etc.). However, this can sometimes lead to increased expenses such as higher labour costs, more fencing to build and maintain, as well as higher management expenses. While feeding the entire herd as one group reduces costs and addresses average herd needs, it can lead to other challenges such as over- and/or under-feeding, which can lead to wasted feed resources [2].

These issues highlight the need to find the balance between meeting a herd's feed/nutritional needs while doing so in a cost-effective manner. Analyzing the chemical composition of feed and forage to understand their nutrient content is another strategy for more optimal cow/calf feed management. This approach not only allows producers to more precisely meet the nutrient requirement of their animals, but also helps producers to decrease expenses or generate extra profits. For example, if a farmer produces high quality forage but his/her herd has

relatively modest nutrient needs, the best decision to bring extra revenue would be sell his high quality forage and buy lower quality forage for feeding his/her herd during the winter [1]. All of the above-mentioned practices demonstrate that proper management of feed costs is crucial to improve profitability in the beef industry.

As noted earlier, the highest feed cost period typically occurs during winter when animals cannot eat fresh forage. In Western Canada, winter usually coincides with the second half of a cow's gestation period (most producers breed their cows in June, July, or August with calving occurring in March, April, or May) when the most critical nutrient needs of dams should be provided. Almost 75% of ruminant fetal growth occurs in late gestation, and therefore nutrient insults at this stage will have the greatest impact on fetal size (i.e., birth weight and body length) [3, 4]. Therefore, modest malnutrition or nutrition restriction during first half of gestation is considered to be of little significance post-natally since the fetus needs relatively little nutrition for growth and development at this stage. Indeed, studies have shown that calves born from dams that were nutritionally restricted during early- to mid-gestation then re-alimented before parturition had similar ratios of weight gain to feed intake compared to fully-fed control groups [5, 6]. However, it is not clear how lower planes of nutrition during early gestation could potentially affect key efficiency and production traits of the progeny animals and this possibility needs additional research.

In addition to adopting optimized feed management strategies to reduce feed costs and increase profitability, genetic selection (for animals with optimal feeding responses) has also been used to improve profitability. In particular, the selection for animals which eat less but produce the same amount of meat have attracted considerable attention from beef producers. One such measure of feed conversion efficiency is called residual feed intake (RFI). The use of RFI in

livestock breeding programs is growing rapidly, especially in the beef industry, since those cattle which are genetically classified as Low-RFI (LRFI, efficient animals) eat less and produce less methane per unit weight gain [7]. However, because RFI selection is relatively new to the beef industry it will be important to investigate if other traits, either positive or negative, may be co-selected along with LRFI. Although both RFI and maternal nutrition during gestation may affect progeny traits, our understanding of the biological mechanisms by which this occurs is poor.

I believe that an improved understanding of RFI, maternal nutrition and progeny traits can be achieved through the help of multiple omics sciences including metabolomics, transcriptomics and epigenomics. Therefore, my thesis is focused on investigating the impact of selection for divergent genetic potential for RFI and maternal nutrition during early- to mid-gestation on progeny bull calf traits using multiple omics technologies. In this chapter, the importance of RFI as a measure of feed efficiency and how it can affect different traits in cattle, are reviewed. In addition, the influence of prenatal malnutrition on beef cattle progeny traits is discussed as well as the application of different "omics" approaches towards assessing progeny traits is reviewed.

1.2. RFI AS A MEASURE OF FEED EFFICIENCY

There are several methods by which feed efficiency is measured in cattle. One such measure is the feed conversion ratio (FCR). FCR is a ratio of feed intake to body weight gain. It is often used as a feed efficiency measurement in the beef industry as it is straightforward to understand, and it is often expressed for a group of cattle in one pen, making it easy to measure. Animals with a high FCR consume higher quantities of feed per kilogram of body weight gain compared to their low FCR counterparts [8]. FCR has significant phenotypic and genetic correlations with feed intake, growth rate, and mature animal size [8]. Since FCR represents a gross measure of feed intake, it cannot distinguish between maintenance and growth requirements of an animal [9]. Therefore,

selection for higher (improved) FCR could result in cattle that grow faster, but which also have a larger mature animal size and higher maintenance and feed requirements [8, 10-13].

RFI is another feed efficiency measure calculated using the difference between an animal's actual feed intake and its expected feed requirements for maintenance and growth over a specific time period [14-17]. RFI was first proposed by Byerly [18] and then by Koch et al. [14]. Koch et al. [14] separated the feed intake of beef cattle into two categories including expected feed intake for a given level of production, as well as the remainder, or residual feed intake. They classified those cattle having greater negative residuals as more efficient [14]. Selecting for LRFI (more efficient) animals is gaining popularity among beef producers and is expected to increase [19], since those cattle which are genetically classified as LRFI eat less and produce less methane per unit weight gain [7]. Differences in RFI appear to be due to differences in energy partitioning, basal metabolism, body composition, as well as activity and energy expenditure [16, 20, 21]. Several studies have indicated that LRFI cattle have reduced feed intake by 9-15% at equal weight and average daily gain (ADG) [16, 22-24]. LRFI cattle also exhibited an improved feed conversion ratio (FCR) by 10-15% at equal weight and ADG, compared to their High-RFI (HRFI) counterparts [16, 22, 24].

Methane is a greenhouse gas, which is produced by ruminant animals during digestion and fermentation [25]. Globally, livestock are responsible for the emission of ~18% of the world's greenhouse gases and therefore reducing their carbon footprint is a key factor in reducing global warming. Several studies have shown that selecting for LRFI cattle is associated with reduced methane production [7, 17, 23]. Compared to HRFI cattle, 25-28% lower methane production in LRFI animals has been reported [7, 17]. Therefore, selection for feed efficiency can favor both the farmer (decreased production costs) and the environment (lower methane). Basarab et al. [7]

estimated that a mature LRFI cow would have a net economic profit of $46/head/year compared to that of HRFI, suggesting that selecting for LRFI could have a significant effect on reducing the costs of production.

Another advantage of RFI over other measures is that RFI is independent of growth characteristics such as body weight (BW) and average daily gain (ADG) [8, 14, 26]. Indeed, a study by Herd et al. [27] showed that selection for LRFI in beef cattle reduced feed intake without compromising body size or growth. Studies in Alberta have found that there appears to be no difference in growth, carcass yield and quality grade between LRFI and HRFI cattle [16, 28]. In a study described by Castro Bulle et al. [29], similar dressed carcass yield, backfat thickness, yield grades, and marbling scores were seen between LRFI and HRFI cattle. Fitzsimons et al. [23] also reported no differences could be detected in muscle depth and fat depth as well as in ultrasonic fat measures between LRFI and HRFI cattle. However, a positive correlation between ultrasound backfat measures and RFI was reported by Basarab et al. [16] and Arthur et al. [22]. Richardson et al. [30] also reported that Angus steers born from LRFI parents had more whole-body protein and less whole-body fat compared to progeny steers of HRFI parents. These conflicting results suggest that further research is needed to understand the relationship between body fat and RFI.

RFI has moderate heritability (h^2 = 0.29-0.46) in cattle, which makes it a good candidate for genetic improvement [16, 22, 28]. When heritability is moderate-to-high, it may provide an opportunity to select for more efficient cattle having lower energy requirements for maintenance [31, 32]. However, the difficulty of measuring individual animal BW and feed intake over a standardized 76-d RFI test as well as the relatively high cost is somewhat prohibitive [16, 33]. Therefore, the development of genetic or metabolomic markers to predict RFI is an attractive alternative which would benefit the beef industry.

1.2.1. Physiological basis for RFI

About 73% of the variation in RFI can be explained by differences in energy expenditures from metabolic processes, body composition, and physical activity, while the other 27% is related to variations in ion transport and proton leakage [19]. More specifically, the proportion of variation in RFI that these processes explain are protein turnover along with tissue metabolism and stress (37%); physical activity (10%); digestibility (10%); heat increment and fermentation (9%); body composition (5%); and feeding patterns (2%) [19]. There are five major physiological processes associated with variation in RFI including: 1) intake of feed, 2) digestion of feed, 3) body composition and metabolism, 4) activity, and 5) thermoregulation. These are described below:

1.2.1.1. Intake of feed

Variation in feed intake is associated with variation in maintenance requirements of ruminants [19]. The more an animal consumes food, the more energy needs to be expended to digest the food. This is, in part, because of an increase in size of the digestive organs and increase in energy expended within the tissues themselves [19]. This energy is known as the heat increment of feeding (HIF), and in ruminants this accounts for as much as 9% of the ingested metabolizable energy (ME) [34]. Several studies have shown a positive correlation (r = 0.08-0.62) of RFI with feeding frequency, feeding duration and eating rate (g/min) [13, 28, 35, 36]. Basarab et al. [37] reported that the higher energy requirement for feeding activities for HRFI crossbred heifers is, in part, due to the higher frequency of daily feeding events (events/day), and the longer duration of head-down time (min/day) than their LRFI counterparts. Other studies also reported that HRFI cattle spend 2-5% more energy in feeding activities than LRFI cattle [20, 37], since they can have 14-22% more daily feeding events than LRFI cattle [13, 28]. Manafiazar et al. [38] also reported that LRFI heifers consumed 5.3% less forage when expressed as kg DM d^{-1} (5.1% less when expressed as a

percentage of body weight) with no negative impact on their BW, back-fat thickness, or ADG compared to HRFI heifers. Given that selection for genetic potential for RFI is associated with variation in feed intake, the more efficient animals (LRFI) could be expected to have less energy expended as HIF.

1.2.1.2. Digestion of feed

Herd and Arthur (2009) reported that feed digestibility accounts for 10% of the genetic variation in RFI. Some studies have shown a negative correlation between RFI and digestibility in cattle [7, 17, 39, 40], indicating that LRFI cattle tend to have greater digestibility [19]. Nkrumah et al. [17] reported that dry matter (DM) and crude protein (CP) digestibility tended to be greater for LRFI as compared to HRFI steers. Basarab et al. [7] also reported that RFI was positively associated with DMI, feeding event duration and bunk attendance. In particular, feeding event duration was negatively correlated with DM ($r = -0.55$) and CP ($r = -0.47$) digestibility, indicating that lower feeding durations were associated with lower DMI and improved digestibility. However, other studies found no significant difference in apparent DM digestibility between HRFI and LRFI cattle [41, 42]. Therefore, further research is needed to reveal the association between feed digestibility and genetic variation in RFI.

1.2.1.3. Body composition and metabolism

Several studies have investigated the relationship between RFI and a variety of carcass quality characteristics. Some have shown that there is no significant association of cattle RFI with muscling characteristics such as loin muscle area and intramuscular fat, at slaughter [27, 28, 43, 44]. On the other hand, some studies from ultrasound scans of subcutaneous fat in cattle suggested that genetic selection for RFI altered body composition [22, 30, 45]. For instance, Arthur et al.

[22] reported that backfat thickness (12th/13th ribs) measured at the end of a post-weaning test for RFI was positively correlated with RFI in beef cattle (r = 0.17). Richardson et al. [30] reported that steer progeny of HRFI parents have less whole-body chemical protein and more whole-body chemical fat than progeny of HRFI parents. It was estimated that these differences accounted for 5% of the genetic variation in RFI. Basarab et al. [37] also reported that LRFI heifers had 3.2% less intramuscular fat, 6.8% less subcutaneous fat, and 6.8% less feeding frequency than HRFI heifers.

Differences in lean muscle and adipose tissue composition may be important for determining RFI. The energy cost for deposition of the same weight of fat tissue is higher than that of lean (protein) tissue [19]. However, there is higher variation of protein turnover in lean muscle compared to adipose tissue. Over time, muscle will consume more energy than an equivalent unit of fat. In other words, efficiencies of nutrient use for fat gain (70-95%) are greater than for lean gain (40-50%) [19]. While the results of Richardson et al. [30] do not directly reflect protein turnover, they imply that LRFI steers possess a more efficient mechanism for protein deposition or a lower rate of protein degradation compared to HRFI steers. Similarly, Cruzen et al. [46] reported that less protein degradation occurs in pigs selected for LRFI compared to those selected for HRFI. This may account for a significant portion of the increased efficiency of nutrient use observed in LRFI animals. In support of this, Lefaucheur et al. [47] reported that LRFI pigs had leaner carcasses with higher muscle content, lower backfat thickness, and lower intramuscular fat content in four muscle tissues including LM (3rd/4th lumbar vertebra level), semimembranosus, biceps femoris, and rhomboideus muscles. In terms of lean muscle content in cattle, two studies have reported negative correlations between RFI and estimated lean content [12], and dissectible

carcass lean content [16]. These results suggest that LRFI animals may potentially have increased retail meat yield over HRFI animals.

1.2.1.4. Activity

Mechanisms associated with differences in animal activity include the work involved in feeding, ruminating, and moving at various speeds. Differences in activity levels are also associated with differences in RFI [20]. Several studies on monogastric species have revealed that variation in RFI is partly a result of variation in activity levels [20, 48-51]. In particular, de Haer et al. [48] reported a positive correlation between the total daily feeding time ($r = 0.64$) and number of visits to a feeding station ($r = 0.51$) with RFI in pigs. Luiting et al. [49] reported that about 80% of the variation in RFI could be related to variations in physical activity in chickens. Variations in activity can also be associated with differences in RFI in cattle. Richardson et al. [52] reported a positive correlation ($r = 0.32$) between RFI and daily pedometer count. It was estimated that these activity differences accounted for 10% of the observed variation in RFI. Herd et al. [20] reported that HRFI cattle had ~ 5% more feed energy intake cost due to their higher level of activity, compared to LRFI cattle. Therefore, these studies show that there is a greater energy cost associated with activities for HRFI compared to LRFI animals.

1.2.1.5. Thermoregulation

In cattle, about 9% of the variation in RFI can be explained by inherent differences in heat production [19]. Those cattle having a higher core body temperature than others under similar conditions (i.e. similar dietary intake) must spend more dietary energy on metabolic heat production at the expense of productivity (i.e. milk or muscle synthesis) [53]. Evaporative heat loss is considered as an energy loss in ruminants, which is mainly regulated by the respiration rate

(RR) [54]. The thermal status of an animal can also be measured via skin (surface) and core (rectal) temperatures (TR), as well as heart rate [54]. For instance, Barea et al. [55] reported that HRFI pigs had greater total heat production, measured in respiration chambers, compared with LRFI pigs, mainly due to their higher physical activity. DiGiacomo et al. [54] reported that HRFI Holstein-Friesian cows had higher skin, neck, and shoulder temperatures as compared to LRFI cows, while the rectal temperature and RR were not influenced by RFI variation. In contrast, Martello et al. [56] reported higher skin temperature in LRFI *Bos indicus* cattle maintained in tropical conditions in Brazil compared to their HRFI counterparts. The authors speculated that the higher skin temperature in the LRFI group may be related to their improved efficiency of thermoregulatory mechanisms since the rectal temperature remained lower in the LRFI group. Clearly, further research is needed to reveal more details regarding the association between energy loss as heat and variation in RFI.

1.2.2. The relationship between RFI and enteric methane

Methane (CH_4) is an important greenhouse gas that contributes to global warming. Methane is produced in ruminants by rumen microbes during digestion and fermentation of their feed [25]. The global warming potential of methane is about 25-fold higher than that of CO_2 [57]. Livestock are responsible for approximately 18% of greenhouse gas emissions [25, 58]. Apart from reducing feed costs, selecting for LRFI cattle can also reduce methane production. Basarab et al. [7] reported that LRFI cattle had 15-25% less methane production compared to HRFI cattle. Nkrumah et al. [17] also reported that LRFI feedlot steers produced lower levels of methane (28%) than HRFI cattle on concentrate diets in the drylot. The higher level of methane production in HRFI cattle is most likely due to the fact that HRFI animals have higher dietary energy intake at equal levels of production compared to LRFI animals [24, 59]. The lower feed intake results in a host-mediated

response in microbial communities (bacteria, ciliate protozoa, fungi, and archaea) in LRFI animals favoring reduced methane production [17, 60]. This was illustrated in several studies where a clear segregation of rumen microbial profiles was seen in cattle with different RFI values. In particular, the ruminal bacterial profiles (but not the total numbers of bacterial cells) were different between steers divergent for RFI when fed both growing [61] and finishing diets [61, 62]. This segregation of ruminal microbial profiles was also seen between forage-fed beef LRFI and HRFI heifers [63]. Therefore, selection for LRFI animals favors a reduced carbon footprint, which appears to be mediated by intrinsic differences in the ruminal microflora of HRFI and LRFI animals.

1.2.3. RFI and its relationship to immune response and stress

Although several studies have been conducted to understand the relationship of RFI with important traits such as metabolism, fertility, and carcass characteristics, little is known about the contribution of the innate immune and stress responses to divergence in RFI, especially in cattle. Indeed, most studies done to date on the effect of RFI on health have focused on pigs. Grubbs et al. [64] reported that LRFI pigs had increased levels of heat shock protein (HSP) 60 and HSP70 in LD muscle and liver, compared to their HRFI counterparts. HSPs have been linked to anti-apoptotic pathways in the mitochondria. They also suggested that LRFI pigs may be less prone to muscular oxidative stress (due to their increased level of anti-oxidant defense pathways), and their liver may have a greater metabolic capacity (pathways leading to reduced oxidative stress, and higher metabolism and cellular repair) compared to their HRFI counterparts. Colpoys [65] also looked at pigs in evaluating the physiological stress response of gilts and divergent selection for RFI by administering adrenocorticotropic hormone (ACTH) and measuring the stress response through cortisol concentrations. Pigs secrete ACTH from the anterior pituitary gland following a stressful event, which then stimulates the secretion of cortisol from the adrenal cortex [65]. The

results showed that LRFI gilts had lower cortisol concentrations in response to the ACTH stimulation compared to HRFI gilts, suggesting that in a production environment LRFI gilts were better able to cope with a physiological stressor than HRFI gilts.

Evidence also suggests that RFI may have an effect on immune function. By analyzing pigs, Rakhshandeh et al. [66] evaluated the impact of lipopolysaccharide (LPS) challenge, a bacterially derived immune system stimulant, from the two RFI lines on the apparent ileal digestibility and apparent fecal digestibility of nutrients. The results showed that immune system stimulation increased the apparent fecal digestibility in LRFI pigs but had no effect on apparent ileal digestibility of nutrients. On the other hand, Azarpajouh et al. [67] reported no differences in sickness behaviors in response to a (LPS) challenge between divergent RFI pigs. A study by Dunkelberger et al. [68] using two RFI lines of pigs infected with porcine reproductive and respiratory syndrome virus (PRRSV) showed that LRFI animals had greater growth under challenge and tended to be less affected by the PRRSV-challenge than HRFI animals. This suggests that LRFI pigs were more robust to the viral challenge. Consolo et al. [69] reported greater lymphocyte count with fewer segmented neutrophils in both LRFI pregnant and non-pregnant heifers compared to their HRFI counterparts. In addition, LRFI non-pregnant heifers showed higher level of immunoglobulins M (IgM) than inefficient calves. Similarly, Herd et al. [70] reported a negative correlation of RFI with white blood cells and lymphocytes count in Angus cattle. Overall, the data suggest that LRFI animals cope with physiological and pathogenic stress better than HRFI animals.

1.2.4. Relationships between RFI and fertility

Even though the reduction of feed costs is a major factor favoring the profitability of beef production, successful reproduction in cow-calf operations is also a key affecting profitability [71].

A study conducted by Smith [72] found that LRFI heifers reached puberty earlier, conceived at an earlier age, and delivered heavier calves compared to their HRFI counterparts. Contrary to this finding, several recent studies have found that LRFI heifers showed delayed sexual maturity [73-75] but no difference was observed in pregnancy or conception rates between LRFI and HRFI heifers [75]. When RFI was adjusted for body fatness (RFIfat), in an attempt to ensure independence with age at puberty in heifers, there were no difference in the percentage of LRFIfat and HRFIfat heifers reaching puberty [35, 37, 44, 73]. In terms of bull fertility, several studies have shown that scrotal circumference (SC) and semen characteristics were not statistically different between LRFI and HRFI bulls. It has also been observed that a greater number of progeny per sire was seen for LRFI bulls in a multi-sire natural mating experiment [76]. In contrast, other studies found that LRFI bulls had smaller scrotal circumference [73], decreased sperm motility [73, 76, 77], decreased progressive sperm motility, increased abundance of tail abnormalities, and delayed sexual maturity [78, 79]. Moreover, when RFI was adjusted for backfat thickness, LRFIfat bulls still had lower sperm motility, progressive motility, as well as a smaller SC [77]. These data suggest that some fertility issues may exist with certain LRFI animals.

In conclusion, the scientific community has reported that variation in RFI is associated with five major physiological processes including: 1) intake of feed, 2) digestion of feed, 3) body composition and metabolism, 4) activity, and 5) thermoregulation. It has also been established that there are clear relationships between RFI and traits such as enteric methane, immune response and fertility. Therefore, the differences in the way high and low RFI cattle use and partition energy might have differential effects upon the offspring of these animals when exposed to contrasting planes of prenatal nutrition. It is important to more clearly define the biological pathways

underlying the biology of RFI in order to predict production responses in different environments that cattle can experience.

1.3. PRENATAL NUTRITION RESTRICTION

Maternal stress, particularly nutritional stress, is considered to be one of the major drivers of any negative consequences arising during the developmental programming of offspring. Research on children conceived prior to and during the Dutch Hunger Winter (between December 1944 and April 1945 in World War II) demonstrated that children exposed to poor nutritional conditions in early gestation had a decline in cognitive function, increased risk of coronary heart disease, and an atherogenic lipid profile, despite having a slightly higher than average birth weight [80, 81]. This study and other similar studies resulted in a proposal by a British epidemiologist, David Barker in 1990. This proposal, called the Barker hypothesis, states that there is a significant association between the occurrence of hypertension and coronary heart disease at middle age in people with premature birth or low birth weight [82]. Since then the idea that prenatal nutrition has long-term impacts on the health of the offspring, has become widely accepted. For the purposes of this discussion, maternal nutrient restriction includes any event which reduces nutrient supply to the fetus during critical developmental stages [83, 84]. Similar to studies conducted on humans, maternal malnutrition during gestation can also affect post-natal growth and development, fertility, and the health of offspring in cattle [3-5, 85].

1.3.1. Effect of timing of prenatal nutrient restriction on growth

The effects of maternal malnutrition on fetal development depends upon timing, level, and/or duration of the period of compromised nutrition [84, 86-88]. Several studies have shown that nutrient insults during the last two-trimesters of a pregnancy can reduce fetal growth and the birth

weight of offspring in sheep [87, 89, 90]. Almost 75% of ruminant fetal growth occurs in late gestation, and therefore nutrient insults at this stage would be expected to have the greatest impact on fetal size (i.e., birth weight and body length) [3, 4]. In a developing fetus, most of the muscle fibre growth (myogenesis) as well as adipocyte growth (adipogenesis) take place during late gestation. In particular, the majority of muscle fibres are established during secondary muscle fibre development (beginning at about the 3rd month of pregnancy and lasting until about 7th or 8th months of pregnancy). Similarly, the development of adipocytes (fat cells) is thought to span the last 5 months of pregnancy and continue after birth in cattle [3]. Therefore, any decrease in birth and body weight because of nutrient insults at these later stages of pregnancy can lead to an impairment of adipogenesis and myogenesis as well as a decrease in muscle fibre growth [3, 4].

In beef cattle production, undernutrition during first half of gestation is considered to be of little significance post-natally since the fetus needs only a limited level of nutrition for growth and development at this stage compared to maternal needs [91-94]. However, critical events such as placental development and organogenesis occur during early pregnancy. In terms of nutrient partitioning in the fetus during gestation, the priority is with essential organs (i.e., heart, liver, lung, brain, kidney, etc.) rather than skeletal muscle. Because no net increase in muscle fibre numbers occurs after birth, nutrient insults during early gestation may change nutrient partitioning to essential organs rather than skeletal muscle [3-5, 85, 95, 96]. However, studies on sheep and cattle have shown that male lamb progeny and calves born from dams that were nutritionally restricted during early- to mid-gestation and then re-alimented before parturition had similar ratios of weight gain to feed intake compared to their normally fed control groups [5, 6, 97]. More recently, Johnson et al. [73] reported that Angus bulls born from dams that were nutritionally restricted during the

first half of gestation and then re-alimented had similar birth weights but were heavier between 10-16 months of age, compared to bulls from normal prenatal diet fed heifers.

1.3.2. Effect of prenatal nutrient restriction on muscle growth

Muscle growth occurs in two ways: 1) muscle hyperplasia/synthesis or myogenesis and 2) muscle hypertrophy. In muscle hyperplasia, the number of muscle fibres is increased while in muscle hypertrophy, the size of muscle fibres is increased [98]. Since both muscle hyperplasia and hypertrophy occur *in utero* in cattle [3], prenatal nutrition can impact muscle growth in multiple ways. With regard to muscle hyperplasia, the vast majority of muscle fibres are formed *in utero* and there is very limited development or increase in muscle fibre numbers during post-natal growth [3]. However, hypertrophy occurs during late gestation and throughout life [3]. Hence, while prenatal nutrition can affect both muscle hyperplasia and hypertrophy, it may have a greater impact on muscle hyperplasia.

As reviewed by Du and coworkers [3], there are two stages of myogenesis in cattle development. During the first wave of myogenesis, which occurs within the first two months of gestation, the primary muscle fibres are formed [3]. Since only a limited number of muscle fibres are established at this stage, nutritional insults during this period likely have a negligible effect on fetal skeletal muscle development [3, 96]. During the second wave of myogenesis, which occurs after the 2^{nd} and up to the 8^{th} month of gestation, the secondary muscle fibres are formed [3]. Since the majority of muscle fibres are established at this stage, nutritional insults during this period appear to negatively affect fetal skeletal muscle development [3, 96]. It is also important to mention that skeletal muscle matures during late gestation, at approximately day 105 for ewes and day 210 for cattle [3]. Hence, nutrient restriction after this stage has no major influence on the number of muscle fibres, although it does impact the size of muscle fibres (Greenwood et al.,

1998). For example, McCoar and coworkers [99] reported that nutrient restriction in late gestation did not affect the number of muscle fibres in lambs but it did reduce the diameter of muscle fibres [99]. Overall, the impact of prenatal nutrient restriction on muscle growth in cattle depends on its timing, with the most negative effect occurring between day 60 to day 210 of gestation, when most muscle hyperplasia occurs.

1.3.3. Effect of prenatal nutrient restriction on carcass quality

Several studies have shown that poor prenatal nutrition is not only associated with reduced growth and body weight of cattle [100], but also negatively impacts carcass quality [6, 101]. For instance, Cafe et al. [100] reported reduced birth weight in calves born to dams that were nutritionally restricted during gestation. Greenwood et al. [101] examined the offspring born from the experiment of Cafe et al. [100] and found that low birth weight calves had smaller carcasses and smaller primal cuts after their slaughter at 30 months of age, compared to high birth weight calves. Long et al. [6] examined the impacts of early- to mid-gestational undernutrition (70% of NRC recommendations) on offspring growth, carcass characteristics, and adipocyte size in beef cattle, compared to a normally fed control group (100% of NRC recommendations). Surprisingly the results showed that there were no significant differences in body weight or organ weights among the two treatment groups. However, yield grade was reduced and semitendinosus weight/hot carcass weight (HCW) tended to be reduced in restricted compared with non-restricted group. In addition, average adipocyte diameter was increased in subcutaneous, mesenteric, and omental adipose tissues. Likewise, perirenal adipose tissue tended to be increased in restricted compared to non-restricted group. This suggests that maternal undernutrition in cattle may cause alterations in carcass quality of the offspring, potentially resulting in greater adiposity and reduced muscle mass. However, Blair et al. [102] reported no differences in HCW, dressing percent, ribeye area,

marbling score, and intramuscular fat between calves born from energy restricted vs. non-restricted cows during mid-gestation. On the other hand, rib backfat and the USDA (the United States Department of Agriculture) yield grade tended to be lower in calves born from energy-restricted dams. These studies show that more research is still needed to evaluate how maternal undernutrition may influence carcass traits in beef cattle offspring.

1.3.4. Effect of prenatal nutrient restriction on health

Several studies have been performed to understand the impacts of prenatal under- or over-nutrition on the health of offspring. Gardner et al. [103] reported that an energy restricted diet (50% restriction in energy) from 110 days of pregnancy until parturition in sheep did not affect the birth weights of lambs but it did reduce glucose tolerance causing insulin resistance in lamb progeny. Nutrient restriction also resulted in increased adiposity in lamb progeny carcasses at 1 year of age, compared to the adiposity measured from non-feed-restricted sheep. In beef cattle, maternal nutrient restriction (~75% of recommended allowance) during early gestation compromised placental angiogenesis, cotyledon weight, and fetal development [104]. Corah et al. [105] reported increased mortality rates of calves born from dams having an energy-restricted diet (65% of their dietary energy requirements) in the last trimester of gestation compared to those born from non-energy-restricted dams (100% of their dietary requirements). While maternal malnutrition during gestation is associated with reduced neonatal viability, enhancing maternal nutrition during gestation may improve the life of progeny animals. In ewes, a study conducted by Budge and coworkers [106] showed higher viability in lambs born from ewes fed 150% of their energy requirements compared to those born from ewes fed 100% of their energy requirements.

Passive immunity, via transfer of maternal antibodies to the fetus through placenta and to offspring through high quality colostrum, is an important factor in determining the health of

offspring [107]. Qi et al. [107] reported that immunization of pregnant mice against *Yersinia pestis* could passively transfer antibodies from mothers via both placenta and colostrum to newborn mice. This passive immunity can persist up to 3 months in newborn mice and as a result can be an effective method against plague epidemics. A study by Hough et al. [108] showed that colostral immunoglobulin G (IgG) concentrations were similar in both Angus cows that consumed just 57% of their needed nutritional requirements in the last trimester of pregnancy and non-feed-restricted cows [108]. However, calves fed colostrum from diet-restricted dams tended to have lower IgG concentrations in serum [108]. Therefore, maternal nutrition is an important factor in determining the immunological health of progeny animals.

1.3.5. Effect of prenatal nutrient restriction on fertility

In terms of the effect of maternal nutrient restriction on the fertility of ruminant offspring, a number of effects have been observed. These include a decreased ovulation rate [109], decreased ovarian weight [6], and a reduced number of postnatal follicles [110] in female progeny. These effects were observed when dams were nutritionally restricted during early- to mid-gestation and then re-alimentated thereafter. In both sheep and cattle, there was no significant effect on reproductive development and adult reproductive function in rams or bull progeny, when ewes or heifers were nutritionally restricted during early- to mid-gestation and re-alimentated thereafter [73, 109, 111]. However, delayed prostate development has been observed in male rat pup progeny, when Wistar rat dams were fed a protein-restricted diet throughout the entire gestational period [112]. In addition, nutrient insults during pregnancy resulted in male lamb progeny with delayed puberty, decreased plasma testosterone concentrations (from birth until 28 weeks of age) and decreased testicular volume (from birth until 35 weeks of age) [113]. Therefore, maternal undernutrition can

affect reproductive development and function in both male and female progeny, although the long-term effects (beyond 35 weeks of age) have not been investigated.

In summary, the availability of nutrients to the dam and the fetus at different stages of gestation can affect the growth and development in the progeny animals. Important production and economic traits in beef cattle such as muscle growth, carcass quality, health, as well as reproductive development and function can be affected. These findings depict the main drawbacks of prenatal undernutrition at different stages of gestation in cattle and highlight that a more thorough understanding of how prenatal nutrition affects post-natal growth and development, and by which mechanisms it does this, is vital to optimize production efficiencies.

1.4. APPLICATION OF MULTI-OMICS APPROACHES IN THE BEEF INDUSTRY

The application of high-throughput technologies has revolutionized biological research by allowing scientists to explore the molecular details of complex biological systems. These high throughput "omics" techniques employ genomics, epigenomics, transcriptomics, proteomics and metabolomics. Genomics enabled by high-throughput, next-generation DNA sequencing has allowed the detailed genetic characterization of thousands of organisms. Epigenomics, which now uses high throughput DNA sequencing technology, is allowing researchers to map DNA methylation hotspots and temporal DNA methylation changes across multiple different cells and tissues. Transcriptomics, as measured via microarrays or, more recently, RNA-Sequencing (RNA-Seq) has allowed thousands of RNA transcripts to be measured and their abundance to be easily compared between cells and tissues. Proteomics has opened the door large-scale protein characterizing and the enumeration of protein abundance changes in cells, tissues and organs. Metabolomics offers a powerful new route to measure hundreds to thousands of metabolites – the downstream products of gene/protein activities – in both biofluids and tissues. Individually,

genomics, transcriptomics, proteomics and metabolomics are called “omics” approaches, together they are called “multi-omics”. Bringing together multiple omics approaches to look at a biological problem offers researchers the opportunity to explore the system, at a molecular level, from many different perspectives. Therefore multi-omics approaches, when combined with advanced bioinformatics techniques offer a more holistic or system-based perspective to explore molecular biology and understand molecular physiology. This section describes and discusses the application of three different omics sciences, metabolomics, transcriptomics and epigenomics, to explore the biological pathways connected with prenatal maternal diet and/or RFI in beef cattle.

1.4.1 Definition of a metabolome and how it is detected and measured

Metabolites are considered as the end products of cellular regulatory processes occurring inside the cell (many of which are guided by the genome) as well as events, exposures or phenomena occurring outside the cell or organism (which are dictated by the environment) [114, 115]. As a result, measuring the metabolome (the complete set of metabolites in a cell, tissue, organ or organism) can help to reveal key interactions between genes and the environment. In other words, metabolomics allows researchers to gain a more complete understanding of an organism’s chemical phenotype [116, 117]. The relationships between the genome, transcriptome, proteome, and metabolome are illustrated in Figure 1.1.

Metabolomics uses advanced analytical chemistry techniques to comprehensively detect and measure hundreds of small molecule metabolites in cells, biofluids and tissues. Metabolomics uses a variety of analytical chemistry techniques including nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) as well as coupled techniques such as liquid chromatography-mass spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS), and inductively coupled plasma-mass spectrometry (ICP-MS). Each technique enables the

detection of specific classes of metabolites with differing levels of sensitivity and accuracy. Therefore, each technique has its own pros and cons.

The application of NMR spectroscopy to our understanding of metabolism and metabolic processes dates back to the mid-1970s [118]. NMR was one of the first technologies to be adopted in the early days of metabolomics and has an exceptional capacity to handle complex metabolite matrices [119, 120]. NMR is a spectroscopic technique that measures and detect perturbations of local magnetic fields around atomic nuclei. For conduction of an NMR experiment, molecules or solutions of molecules are first placed in a strong magnetic field to orient each of the molecule's nuclear magnetic dipoles (found within each atom's nucleus). Then a weak oscillating magnetic field is passed through the sample. This radio-frequency magnetic oscillation leads certain excited nuclei within the molecule to resonate and absorb the magnetic energy and to emit a detectable electromagnetic signal with a frequency characteristic of the magnetic field at the nucleus (Figure 1.2). The spectral emission peaks that are detected by NMR are called resonances and their position or frequency is called the chemical shift. The chemical shift of a given NMR resonance is affected by various chemical or molecular properties such as electronegativity, proximity to charges or chemical bonding. The intensity of NMR peaks is directly proportional to the number atoms at a given chemical shift, and can be used to characterize chemical structures of molecules very precisely.

NMR has a number of unique benefits over most analytical chemistry techniques such as MS or other kinds of spectroscopies for metabolomics applications. In particular, NMR is nondestructive, nonbiased, easily quantifiable, very reproducible, requires little or no compound separation, permits the identification of novel compounds, and does not require any chemical derivatization [121]. A significant advantage of NMR over GC-MS or LC-MS is the fact that it is

capable of detecting compounds that are less tractable to MS techniques, such as certain sugars, amines, volatile compounds (e.g., acetone, methanol and ethanol) and relatively nonreactive compounds [121]. Although NMR is often chosen for its reliability and utility for absolute quantitation of metabolites, its main disadvantage is its poor sensitivity. Typically, NMR has a lower limit of detection of approximately 1-5 μM and a requirement of relatively large sample sizes (~500 μl) [121]. This makes NMR 100-1000 times less sensitive than other analytical chemistry techniques such as mass spectrometry. The sensitivity (limit of detection) differences among different analytical platforms is illustrated in Figure 1.3.

Analyzing NMR spectra for metabolomics applications requires advanced software such as the Chenomx NMR suite (Chenomx Inc., Canada) or Bruker TopSpin (Bruker, Germany) to perform spectral deconvolution (Figure 1.4). Spectral deconvolution involves extracting or identifying individual spectra of pure compounds from the more complex spectra of a mixture of compounds. As shown in Figure 1.4, the NMR spectrum of a biological sample that contains a number of different metabolites, is actually the sum of individual NMR spectra for each of the pure metabolites in the mixture. As opposed to the concept of a single peak for a single compound in MS techniques, multiple NMR peaks are normally associated with a single compound (Figure 1.4). This fact greatly reduces the problem of spectral redundancy or spectral overlap in NMR spectra of chemical mixtures. This means that with NMR it is unlikely that any two compounds will have identical numbers of peaks with identical chemical shifts or peak intensities. In an NMR spectrum, by properly matching and fitting the observed peaks in the chemical mixture to a library of pure reference compound peaks (which have been calibrated to an internal concentration standard), it is possible to simultaneously identify and quantify dozens of compounds at once [121,

122]. Typically, an NMR-based metabolomic analysis can measure between 35-60 different metabolites in serum with concentrations down to 1-2 □M.

Another commonly used technique in metabolomics is GC-MS. GC-MS has played a key role in our understanding of metabolism and metabolic processes for more than 50 years [123-125]. In GC-MS the analytes of interest are first chemically derivatized, then volatilized (put into the gas phase by heating) and then separated in a long (30 m) GC column which is a thin tube lined with hydrophobic liner and filled with a separation gas (usually helium) maintained at high temperatures (up to 300 °C). The derivatized molecules interact with the GC column and then elute (come off) from the column at different times (called the retention time [RT]). A typical GC run last for 30-45 min, with compounds eluting every few seconds. The derivatized molecules reach the mass spectrometer (which is usually a single quadrupole (Q) or time-of-flight (TOF) detector) where each molecule is broken into ionized fragments and their mass-to-charge (m/z) ratio measured by the mass analyzer. The combination of the time taken by the analyte to travel the GC column (called the retention index [RI]) and the mass fragmentation pattern acquired from the mass spectrometer allows the compounds to be identified and quantified [126] (Figure 1.5).

Similar to NMR, GC-MS can be used to identify and quantify a wide range of metabolites, including organic acids, amino acids, nucleic acids, sugars, amines and alcohols [127-129]. One of the great advantages of GC-MS over NMR is the fact that GC-MS is a very sensitive technique, with a lower limit of detection at the high nanomolar level. It is also a very reproducible, very precise and a highly standardized technology [128]. However, unlike NMR, GC-MS is a destructive process, requiring chemical fragmentation, chemical derivatization and separation of analytes prior their detection and quantification. This chemical derivatization step can introduce a

number of artifacts and reaction side products that can result in false positive compound identification, complicating GC-MS analyses [128].

Analyzing GC-MS spectra requires specialized spectral deconvolution tools such as the Automated Mass Spectral Deconvolution and Identification System (AMDIS), which is produced by the National Institute of Standards and Technology (USA) or Agilent MSD ChemStation (Agilent Technologies Inc., Germany). A GC-MS spectrum or total ion chromatogram for a biological sample typically consists of dozens to hundreds of sharp peaks (corresponding to ion counts) covering an elution time of approximately 30-45 min. Each peak may consist of one or more electron ionization (EI) mass spectra arising from one or more compounds (Figure 1.6). For metabolite identification, each EI spectrum extracted from the total ion chromatogram is compared to spectral reference libraries (e.g. the National Institute of Standards and Technology or the NIST library) containing the EI spectra of thousands of pure, derivatized and authenticated compounds. This spectral comparison allows one to identify and quantify approximately 75-100 different metabolites in biofluids such as serum or urine.

The most commonly used analytical method in metabolomics is LC-MS. The application of LC-MS to metabolomics studies dates back to the early 1980s [130, 131]. The basic principles of LC-MS are very similar to GC-MS. Compounds are separated via a chromatographic column followed by detection with a mass analyzer and finally spectral deconvolution of the total ion chromatogram [121] (Figure 1.7). However, LC-MS differs from GC-MS in several ways including [121]: 1) analytes are not (normally) chemically modified prior to separation in LC-MS; 2) separation of analytes is performed in a liquid phase (over a C18 alkane column) under high-pressure or ultra-high-pressure conditions in a liquid chromatography system instead of in the gas phase under low-pressure conditions; 3) soft ionization methods such as electrospray ionization

(ESI) or atmospheric pressure chemical ionization (APCI) are used in LC-MS while a hard ionization method called electron ionization (EI) is used in GC-MS; 4) in LC-MS the parent ions (M+H ion) are sometimes further fragmented through collision-induced decay to produce product ion fragments or spectral signatures (through the addition of another mass spectrometer called an 'MS/MS' or tandem mass system); and 5) the mass detectors are much more varied (and sensitive) in LC-MS and may include single quadrupoles (Q), triple quadrupoles (QqQ), linear ion traps, OrbiTraps, TOF (time of flight) and Fourier transform ion-cyclotron resonance systems (FT-ICR) as opposed to simple quadrupole or TOF detectors in GC-MS. These differences result in a number of advantages in LC-MS methods over GC-MS methods [128, 132]. In particular, LC-MS instruments are typically 10-1000 times more sensitive than GC-MS or NMR techniques (Figure 1.3). They can measure masses much more accurately over a much larger mass range (50-2000 Da) than GC-MS (50-800 Da). Furthermore, because there is usually no chemical derivatization, LC-MS instruments do not produce chemical artifacts or create hard-to-identify side-products. However, unlike NMR and GC-MS, LC-MS is not particularly reproducible and is usually very poor at separating polar molecules (which constitute the majority of metabolites in biofluids) relative to nonpolar molecules [128, 132]. Furthermore, because LC-MS methods use a wide variety of column types and mobile-phase combinations, it is almost impossible to create standardized retention times or retention indices (as is commonly done with GC-MS), making it almost impossible to identify metabolites using retention times or retention indices.

As a result, identification of metabolites by LC-MS is not as straightforward as GC-MS. Compound identification by LC-MS can be done in one of two ways. One is based on matching MS/MS fragment patterns to spectral libraries (similar to what is done with GC-MS) and the other is based on accurately measuring parent ion masses to determine their molecular formula and then

determine their identity via comparisons to compound databases with known molecular weights and known molecular formulas. For LC-MS spectra matching, instrument- or condition-specific MS/MS product ion fragment libraries are required. Some of these libraries are commercial MS/MS spectral databases specifically developed for drugs, natural products, pesticides and metabolites for a variety of single quadrupole, triple quadrupole and ion trap instruments [128]. There are also a number of freely accessible electronic MS/MS spectral databases such as those found in the HMDB [133-136], MoNA (https://mona.fiehnlab.ucdavis.edu/) and NIST (https://www.nist.gov/).

Most of the metabolomic methods described so far are designed to detect organic molecules. But for the detection of metal ions or other inorganic compounds found in the metabolome, ICP-MS (Inductively Coupled Plasma-Mass Spectrometry) is a powerful and sensitive approach for measuring metal ions and other trace elements in biological and inorganic samples [137]. ICP-MS can also be used to detect different isotopes of the same element making it a versatile tool for isotopic labeling and isotopic analysis for geological or geographic analyses. In the ICP-MS process [138], liquid samples are first vaporized using a nebulizer and then introduced into a hot argon (carrier gas) plasma consisting of electrons and positively charged argon ions. Next, the plasma's high temperature (6000-10000 ^{o}C) splits the sample into individual atoms (atomization). The plasma also ionizes these atoms – causing the atoms to lose electrons and become positively charged ions (anions are not detected by ICP-MS). The positive ion beam enters a quadrupole mass analyzer where the ions are separated according to their mass/charge (m/z) ratio. After passing the quadrupole, the ions hit a special mass detector that detects ions at the higher and lower mass range (Figure 1.8).

Some of the advantages of ICP-MS include its multi-elemental capability (it can detect almost all elements in the periodic table), its ability to perform rapid isotopic analysis, its good precision, a large linear dynamic range (over several orders of magnitude), the availability of simple spectra, and high sensitivity (low detection limits down to concentrations of subfemtograms per gram of matrix solution) [139]. With such advantages, ICP-MS has become the gold standard for identifying and quantifying different elements in both biological and inorganic samples. However, ICP-MS is prone to interference caused by sample matrix components due to polyatomic ions having the same nominal mass as the analyte. In addition, clogging of the ICP-MS inlet is another challenge requiring that operators keep the concentration of dissolved samples very low. Hence, ICP-MS requires sample pre-treatment and good calibration curves for each analyte. Nonetheless, with the appropriate method in place, these limitations can be largely minimized, allowing ICP-MS to be used for the analysis of almost all sample types [139].

Analyzing ICP-MS spectra requires a carefully constructed calibration curve for each metal ion or analyte. This can be done by preparing individual calibration standard dilutions (external calibration) for each analyte, covering the lowest to highest concentrations that can be detected by ICP-MS. Each analyte is compared to its own calibration curve for quantification. An internal standard must be combined with external calibration in order to compensate for matrix effects (the effects of the surrounding material containing the sample of interest) and instrument drift. As a result, the same amount of one or more internal standard elements is added to all measured solutions (blanks, calibration standards, quality control standards, unknown samples, etc.). The response from the internal standard element is expected to be the same throughout the assay, thereby any variation is assumed to be derived from either matrix effects or instrument drift. Typically, a mathematical correction factor is calculated from the relative internal standard

response which is then applied to the analytes to correct for both matrix and drift effects. ICP-MS methods can typically measure 30-35 metal ions down to the femtogram level.

1.4.1.1. Application of metabolomics in the beef industry associated with RFI and maternal nutrition

Metabolomics have been widely used in cattle studies to detect disease [140-142], to assess physiological responses to different diets [143, 144], to evaluate carcass quality [145-147], to assess fertility [148], milk quality [149, 150], and even characterize the near-complete chemical composition of cow's milk [151]. Metabolomics has also been used to predict RFI [13, 23, 152] and to explore the molecular physiological or biochemical underpinnings of lean tissue or muscle content in livestock. For example, Lawrence et al. [42] reported a positive correlation between muscularity and concentrations of the metabolite, creatinine, in blood. Creatinine is produced via creatinine phosphate and typically is increased due to muscle breakdown. In sheep, blood creatinine concentrations have been shown to have a negative correlation with fat depth and a positive correlation with muscle mass [153, 154]. Urea concentrations in blood also showed a negative correlation with lean growth of sheep [153, 154]. Moreover, in the blood of weaned Angus steers, urea concentrations tended to have a positive association with RFI [155].

In terms of RFI, Fitzsimons et al. [23] reported higher concentrations of glucose and urea and lower concentrations of creatinine in plasma of HRFI vs. LRFI heifers. Karisa et al. [152] also reported higher concentrations of creatine and carnitine in the plasma of HRFI beef cattle. Similarly, Kelly et al. [13] reported a higher concentration of beta-hydroxybutyrate in the plasma of HRFI beef heifers. Therefore, there is clear evidence that a number of plasma metabolites differ between LRFI and HRFI animals and that metabolomics offers a good opportunity to detect additional markers of RFI propensity.

While a good deal of metabolomic data has been collected and published on RFI, there is very little published large scale metabolomic data on maternal prenatal diet restriction and their effects on progeny. Indeed, a literature review conducted by us revealed only a few studies that looked at metabolite levels or metabolite changes in the progeny of animals fed with restricted or enhanced diets. For example, Ford et al. [97] reported that the male offspring of ewes experiencing nutrient restriction during early- to mid-gestation and re-alimented afterwards had higher blood glucose levels than those of a control group fed with normal nutrition. In the second study we found, Zambrano et al. [156] reported that the male offspring of rats having a protein-restricted diet during pregnancy had higher levels of cholesterol and triglycerides than the non-restricted group. They also found that cholesterol and triglycerides did not differ between groups for the female progeny. Several studies have also investigated the potential impact of maternal nutrition on the nutrient composition of amniotic fluid (AF) as its composition reflects both maternal health and fetal status. In particular, dietary glucose restriction in pregnant rats resulted in reduced concentrations of methionine and phenylalanine in AF [157]. Similarly, nutrient restriction in pregnant ewes significantly reduced total amino acid and polyamine concentrations in AF [158]. On the other hand, an increased supply of carbohydrates in the diet of pregnant rats significantly increased glucose but also decreased uric acid in AF [159]. Overall, it is clear that the prenatal diet can significantly affect the concentration of a number of metabolites in AF, which possibly results in alterations in the metabolic profile of both the dam and its offspring.

1.4.2. Definition of a transcriptome and how it is detected and measured

While metabolomics allows one to explore metabolic changes associated with physiological perturbations or phenotypic differences, transcriptomics allows one to explore mRNA changes associated with these differences or changes. The genetic information content of an organism is

encoded in its DNA (the genome) and expressed through mRNA transcription (the transcriptome). The transcriptome is the complete set of all RNA molecules in a cell or a population of cells for a specific developmental stage or physiological condition (Wang et al., 2009). Transcriptomics examines the expression level of RNAs in a given cell population and has been widely used in livestock studies [21, 160-162].

Transcriptomics uses advanced molecular biology techniques to comprehensively measure expression of dozens to thousands of transcripts in cells, biofluids and tissues. Such techniques include microarrays or gene-chips, RNA-Seq, and quantitative polymerase chain reaction (qPCR). Each technique has its own pros and cons.

The development of microarray technologies has allowed the measurement of the expression of thousands of genes simultaneously [163, 164]. In this technique, case and control samples of isolated mRNA are labeled with Cy3 and Cy5 fluorescence dyes (via reverse transcriptase PCR) and then the fluorescently labeled complementary DNA (cDNA) "transcripts" produced by the PCR reaction are hybridized to a set of predetermined short (25-70 bp) gene-specific subsequences (cDNA probes) that have been attached to a specially treated glass chip. After the hybridization step is completed (and after suitable washing and fixing steps) the level of gene expression is compared between case and control samples using a specially designed fluorescent microscope scanner, which captures fluorescence signals emitted by the Cy3 or Cy5 labeled DNA strands from the sample that hybridized to the chip DNA probes [163] (Figure 1.9). The more hybridization of the target RNA to a given probe, the more signal. Microarrays offer a number of unique advantages for transcriptome profiling. In particular, they offer a low-cost, high throughput, semi-quantitative method for measuring transcript abundance. They do not require radioactivity (they use fluorescence), they use many advanced micro and nano-scale technologies,

and they can simultaneously measure the expression of thousands of genes. In addition, kit systems and commercial suppliers make microarrays very easy to use [164-166]. On the other hand, microarrays are prone to technical variations (i.e. cross-hybridization artifacts) and have limited ability to accurately quantify very weakly and very strongly expressed genes (i.e., they have a limited dynamic range) [167].

Unlike microarrays, RNA-Seq uses next-generation sequencing (NGS) allowing genome-wide quantification of gene expression in a given cell population [167, 168]. Recent advances in the RNA-Seq workflow, from sample preparation to the operation of sequencing platforms as well as data analysis, has enabled deep transcriptome profiling and a very detailed elucidation of many different biological conditions. The standard RNA-Seq procedure includes isolating RNA from a given sample, converting it to complementary DNA (cDNA), preparing the sequencing library, and sequencing it on an NGS platform [165, 167] (Figure 1.10). Unlike the microarray technique, which is a hybridization-based approach, RNA-Seq is not limited to detecting transcripts that correspond to fully sequenced organisms. RNA-Seq can show the exact location of genomic boundaries, to a single-base resolution. Unlike microarrays, RNA-Seq also has a large dynamic range (it does not have an upper limit for quantification) and it can be used to detect and quantify very weakly expressed genes, although a lot of sequencing is needed to get an accurate quantification of weakly expressed transcripts [167]. RNA-Seq also has much better reproducibility than most microarray platforms [167]. RNA-Seq not only provides a more detailed and quantitative view of RNA abundance than microarrays, but it also identifies alternatively spliced genes, and can detect allele-specific expression if polymorphisms are present in the transcript [165]. In addition to supporting the investigation and enumeration of mRNA transcripts, RNA-Seq can be applied to quantify various populations of RNA, including total RNA, pre-

mRNA, and noncoding RNA (i.e. rRNA, microRNA, and long ncRNA) [165]. However, bioinformatic challenges during data analysis and the high costs of large-scale DNA sequencing are the main drawbacks of RNA-Seq. These issues have limited the widespread application of this technique by the scientific community [167].

qPCR is considered as the most accurate method for measuring the mRNA abundance in cells or tissues [169]. A typical qPCR experiment consists of isolating RNA from a given sample, converting it to cDNA, and then amplifying the sequence from the gene of interest on a qPCR platform (Figure 1.11). There are two types of qPCR typically used by researchers: 1) standard qPCR and 2) high throughput qPCR [169, 170]. A standard qPCR permits the detection and quantification of one gene for up to 384 different samples (in a 384 well plate) or a handful of genes for a smaller number of samples. However, a new generation of qPCR has emerged that permits higher throughput qPCR measurements. In particular, the BioMark HD system (Fluidigm Corporation, San Francisco, USA), uses a microfluidic controller to permit the detection and quantification of 96 genes for 96 samples (thousands of reactions simultaneously in a single run). This dramatically increases the scale and coverage over standard qPCR. The main advantages of qPCR over other transcriptomics techniques are its high accuracy, its excellent reproducibility, and low cost [169, 170]. On the other hand, qPCR is prone to technical errors produced during sample preparation and workup, although there are not as many steps as in preparation for microarrays or RNA-Seq. In addition, qPCR requires that users must design a new primer pair for each gene of interest or obtain the primer pair sequence for each gene from the literature. This manual design step is not required for microarray or RNA-Seq techniques [169, 170]. The high accuracy and reproducibility of qPCR over other transcriptomics techniques has led qPCR to be

the method of choice to confirm the results of other techniques such as microarray and RNA-Seq. Hence, qPCR is considered as the gold standard for measuring gene expression.

1.4.2.1. Application of transcriptomics in the beef industry for characterizing RFI and maternal nutrition

Several studies have been conducted to measure gene expression in cattle, especially in terms of evaluating the impact of RFI and maternal nutrition on gene expression. For instance, gene expression differences in liver and muscle of beef cattle were found to be associated with a divergence in RFI as well as with maternal diet [21, 160-162]. A study performed by Al-Husseini et al. [160] showed that Glutathione S-transferase mu 1 [*GSTM1*] had higher levels of expression in the liver of HRFI Angus steers than in LRFI steers. Kelly et al. [21] reported higher mRNA abundance of Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) [*PPARGC1A*] in the *Longissimus dorsi* (LD) muscle of LRFI compared to HRFI cattle. In a study reported by Paradis et al. [162], five genes involved with inflammatory processes, including hemoglobin β [*HBB*], myxovirus resistance 1 interferon-inducible protein p78 [*MX1*], ISG15 ubiquitin-like modifier [*ISG15*], hect domain and RLD 6 [*HERC6*], and interferon-induced protein 44 [*IFI44*] showed different expression patterns in the liver between HRFI and LRFI heifers.

Maternal diet effects on gene expression have also been explored. A study by Paradis et al. [161] found higher levels of expression of insulin like growth factor 1 [*IGF1*], insulin like growth factor 1 receptor [*IGF1R*], insulin like growth factor 2 receptor [*IGF2R*], insulin receptor [*INSR*], myogenic differentiation 1 [*MYOD1*], myogenin [*MYOG*], and peroxisome proliferator activated receptor gamma [*PPARG*] in the LD muscle of fetal calves from dams exposed to a restricted diet (85% of total metabolizable energy requirements) during mid- to late-gestation compared to a high nutrient diet (140% of total metabolizable energy requirements). Micke et al. [171] reported that

feeding a high protein diet (240% of recommended daily crude protein intake) to beef heifers during the first trimester of gestation was associated with higher expression of leptin [*LEP*] and *IGF1* in the perirenal adipose tissue of both male and female beef cattle offspring at 680 days of age. Furthermore, lower expression of *IGF1R* was seen in the subcutaneous adipose tissue of all progeny between 552 and 657 days of age. These data provide evidence of detectable changes in gene expression in both HRFI and LRFI animals as well as in the progeny arising from maternal dietary differences.

1.4.3. Definition of the epigenome and how it is detected and measured

While transcriptomics offers one route to detect genetic changes or environmentally induced changes in the genome, epigenomics offers yet another route to explore environmentally induced changes to the genome. The epigenome is made up of chemical compounds and proteins which can attach to the DNA strand or histones and direct the genome what to do, such as turning genes on/off. Epigenetic effects are known to be involved in cell growth, the immune response, and in controlling the production of proteins in particular cells [172-174]. Once epigenomic compounds attach to a given DNA strand or to specific histones and modify their function, they are said to have "marked" the genome. These marks do not alter the sequence of the DNA but change the way cells use the DNA's instructions. These marks can be passed on from cell to cell as cells divide and from one generation to the next. These modifications occur naturally in cells and can be altered in response to environmental exposures or disease.

Epigenetics is defined as the study of heritable changes in gene expression, resulting from alterations in chromatin structure but not alterations in the DNA sequence [175, 176]. In contrast to epigenetics, epigenomics pertains to the analysis of epigenetic changes across many genes in a cell or across the entire genome in an organism. Examples of epigenetic mechanisms that produce

such changes are DNA methylation as well as histone modification, each of which alters how genes are expressed without altering the underlying DNA sequence [172]. These epigenetic processes control the timing and intensity of gene expression once cells are differentiated [177, 178].

DNA methylation can directly affect the DNA in a genome, wherein proteins called DNA methyltransferases (Dnmts) attach methyl groups from S-adenyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5-methylcytosine (5mC) in specific DNA locations. These locations are mostly at CpG sites where cytosines precede a guanine nucleotide. These modifications can effectively turn genes on or off [173, 179]. In contrast to DNA methylation which affects DNA directly, histone modification, affects DNA indirectly [174]. In all eukaryotic cells, chromosomal DNA is wrapped around histone proteins, thereby forming spool-like structures. These protein spools enable the long DNA molecule to be wound up neatly into chromosomes inside the cell nucleus. In histone modification, histone proteins can attach a variety of chemical tags (i.e. methyl, acetyl, or phosphoryl group) to their amino acid side chains. As a result, specific proteins can identify these chemical tags and determine whether that specific DNA region should be used or ignored during transcription [174].

Genomic imprinting is an epigenetic phenomenon that causes genes to be expressed in a parent-of-origin-specific manner [180-183]. The genome is passed from cell to cell when cells divide and from parents to their offspring through sexual reproduction. Sexually reproducing organisms inherit two copies of their genes, one from their father and the other from their mother. In cells, usually both copies of each gene are “turned on” or active. However, in some cases, only one copy is “turned on”. Which copy is active depends on the parent-of-origin. Some genes are normally “turned on” only when they are inherited from an individual’s mother; others are active only when they are inherited from an individual’s father. This phenomenon is known as genomic

imprinting (Figure 1.12). These epigenetic DNA marks are made ("imprinted") in the germline (egg or sperm) cells of the parents during meiosis and are maintained through mitotic cell divisions in the somatic cells [184]. Much of the epigenome is reset during meiosis when parents pass their genomes to their offspring [185]. However, in certain cases some of the DNA marks existing in the DNA of eggs and sperm may be passed on to the next generation. When cells divide, often much of the epigenome is passed on to the next generation of cells, helping the cells remain specialized. An individual organism's epigenome can also be changed throughout its lifetime. This usually occurs as a result of changes in lifestyle and environmental factors (i.e., diet, stress, toxins, particulate air pollutant matter and infectious disease), which can expose an organism to pressures that prompt the appearance or alteration of epigenomic marks [179, 186]. Some of these epigenomic marks can be harmful while others can be quite helpful. However, the ability of the epigenome to adjust to these life pressures appears to be required for normal health and development in most organisms [187].

Epigenomics uses advanced molecular biology techniques to comprehensively detect epigenomic modifications in DNA or in histones. The most common techniques currently used for detecting histone modifications are chromatin immunoprecipitation (ChIP)-based technologies [188], including ChIP-Seq, ChIP-qPCR and ChIP-on-chip. On the other hand, whole genome bisulfite sequencing (WGBS) is used to detect DNA methylation patterns in genomic sequences [189, 190].

ChIP is an antibody-based, high-throughput technique primarily used to determine how transcription factors and other chromatin-associated proteins interact with DNA [188]. A typical ChIP procedure consists of cross-linking a protein of interest to target DNA sites in cells, extracting and shearing chromatin by sonicating, adding specific antibodies to immunoprecipitate the target

protein, unlinking the protein from the target DNA, followed by purifying the DNA [188]. Following the completion of a ChIP experiment, protein-DNA interactions can be analyzed using multiple downstream applications, including qPCR (ChIP-qPCR), microarray (ChIP-on-chip), and sequencing (ChIP-Seq) [188] (Figure 1.13). The advantages and disadvantages of high-throughput sequencing, qPCR, and microarray platforms have already been mentioned. ChIP-qPCR can be utilized to detect the interaction between specific transcription factor/histone modification and known genomic binding sites amplified via qPCR. Therefore, ChIP-qPCR benefits those studies focusing on specific genes and potential regulatory regions under different experimental conditions, utilizing low-cost qPCR. Similar to ChIP-qPCR, ChIP-on-chip methods can be utilized to detect the interaction between specific transcription factor/histone modifications and known genomic binding sites, but at a much larger scale covering thousands of these interactions simultaneously. Using ChIP in combination with a microarray, hundreds of DNA sites that interact with specific transcription factors or DNA binding proteins can be analyzed simultaneously. This information can then be linked to the representative gene clusters, to further understand gene expression and regulation. Using ChIP coupled with high-throughput DNA sequencing (ChIP-seq) allows researchers to rapidly identify protein-DNA interactions in a genome-wide manner. In other words, ChIP-seq can be used to accurately map global binding sites for any protein of interest, providing higher resolution and richer or more accurate information, compared to ChIP-qPCR or ChIP-on-chip.

While ChIP techniques are used to identify protein-DNA interactions in the epigenome, specialized DNA sequencing is used to identify methylation patterns in the epigenome. In particular, whole genome bisulfite sequencing (WGBS) is used in the high-throughput genome-wide analysis of DNA methylation status. WGBS involves treating the DNA with sodium bisulfite

before sequencing [189, 190]. Bisulfite treatment of DNA leads to the conversion of unmodified cytosine (C) to uracil (U) while maintaining 5-methyl cytosine (5mC) unchanged as C [189]. After sequencing, the unmethylated cytosines appear as thymines [189]. The standard WGBS procedure includes isolating the genomic DNA from a given sample, modifying the DNA with bisulfite treatment, preparing the sequencing library, and sequencing it on an NGS platform [189] (Figure 1.14). The resulting DNA sequences are then mapped to the reference genome to measure the methylation status of CpG dinucleotides based on mismatches resulting from the conversion of unmethylated C into U. Since WGBS uses NGS technology, its pros and cons are similar to those mentioned for RNA-Seq. However, the main limitation of WGBS comes from the bisulfite modification technique, as incomplete conversion of C to U serves as a source of false positives [189, 190]. In addition, bisulfite treatment can also cause DNA degradation and requires an additional purification step to remove the sodium bisulfite [189, 190].

1.4.3.1. Application of epigenomics in the beef industry associated with RFI and maternal nutrition

While only one study on beef cattle has focused on association between RFI and epigenomic marks, several studies have been conducted to reveal epigenomic modifications as a result of maternal nutrient changes in cattle. In terms of RFI, a recent study conducted by Rocha et al. [191] identified 1493 differentially methylated cytosines (DMCs) and 279 differentially methylated regions (DMRs) in the hepatic tissue of Nelore cattle exhibiting extremes in RFI. In cattle, it has been found that maternal nutrient status can cause epigenetic modifications to the genome of the developing fetus, which potentially can impact the growth, feeding efficiency, fertility and health of future generations [176]. Lan et al. [192] reported that the methylation levels of CpG islands of both H19 fetal liver mRNA [*H19*] and *IGF2R* were higher in the fetuses of pregnant sheep fed

alfalfa haylage (fiber) and dried corn distiller's grains (fiber plus protein plus fat) than those fed only corn (starch). Wang et al. [193] reported that feeding a high starch (corn silage) diet to bovine dams during late gestation increased the expression of the imprinted genes, *IGF2R*, *H19*, maternally expressed 8 [*MEG8*] as well as the DNA methyltransferase gene, DNA methyltransferase 3 alpha [*DNMT3A*], in the LD muscle of progeny male calves compared to low starch (haylage) group. A study reported by Crouse et al. [194] also demonstrated that maternal nutrition during the first 50 days of gestation could alter the expression of histone and histone modifying genes in bovine fetal liver. In particular, nine histone genes were upregulated in fetal liver of moderately nutrient restricted (RES) heifers compared with control (CON) heifers. The upregulated genes include members of the histone H1, H2A, H2B, and H4 families [194]. In addition, 13 histone-modifying transcripts were also differentially expressed. These include genes associated with methylation, phosphorylation acetylation, deacetylation, and ubiquitination. More specifically, these genes were histone deacetylase 10 [*HDAC10*] and the histone deacetylase complex gene, co-repressor interacting with RBPJ 1 [*CIR1*] which showed greater expression in the liver of RES fetuses compared with CON fetuses. The SET nuclear proto-oncogene [*SET*], which is involved in preventing H4 lysine acetylation, was the only gene associated with histone modifications that showed lower expression in liver of RES fetuses. These data suggest that maternal nutrient insults during the early gestation period (i.e. the first 50 days) initiate developmental programming via epigenetic alteration of the fetal genome in beef cattle [194]. Collectively, these studies provide substantial evidence of detectable epigenomic marks in both HRFI and LRFI cattle as well as in offspring arising from maternal dietary differences.

1.4.4. Application of integration of multiple omics sciences in the beef industry

In addition to these "pure" metabolomic, transcriptomic and epigenomic studies, several groups have combined multiple "omics" methods to explore how genetic changes alter the metabolism or metabolome in cattle. In particular, genome-wide association studies (GWAS) have been performed to look for significant metabolomic changes arising from specific single nucleotide polymorphisms (SNP) in cattle [147, 195, 196]. In particular, Weikard et al. [147], as well as Widmann et al. [195] reported that a SNP (I442M) belonging to the non-SMC condensin I complex subunit G [*NCAPG*], is associated with higher arginine levels in plasma. *NCAPG* is a substantial modulator of pre- and/or post-natal growth in cattle. Weikard et al. [147] also reported that a polymorphism in the 204X allele of myostatin [*MSTN*] (previously known as *GDF8*) is associated with higher concentrations of carnitine in the plasma of cattle. Lu et al. [196] reported that higher concentrations of choline and carnitine, as well as lower concentrations of citrate, creatine or phosphocreatine, glycerol-phosphocholine, mannose-like sugar, acetyl sugar phosphate, uridine diphosphate (UDP)-related sugar, and orotic acid in the milk cows with the *DGAT1* KK genotype compared to those with the *DGAT1* AA genotype. This suggests that cows differing in certain *DGAT1* polymorphisms can differ in their milk metabolome. Overall, utilizing multiple "omics" platforms simultaneously in an experiment and then studying the association among the obtained data can help to understand how genetic changes alter the metabolism or metabolome in cattle, and how these genetic and metabolomic variations lead to differences in important production phenotypes.

1.5. CONCLUDING REMARKS

Maximizing the efficiency and growth potential of beef cattle requires not only genetic selection (i.e., selection for feed efficiency, RFI or ADG) but also providing adequate nutrition throughout

all stages of gestation, growth and development of cattle fetuses. Although both RFI and maternal nutrition during gestation can affect progeny traits, our understanding of the biological mechanisms by which this occurs is poor. This thesis is aimed at investigating the impact of RFI selection and maternal nutrition during early- to mid-gestation, on bull calf traits such as muscle development and health. This investigation has been performed using multiple "omics" approaches including metabolomics, transcriptomics, and epigenomics. This thesis is also aimed at identifying candidate biomarker genes and metabolites for feed efficiency, immunity and/or muscle development. The discovery of biomarker genes and/or metabolites associated with these traits of interest will allow producers to select for efficient animals without having to extensively (and expensively) phenotype them, which could be of great benefit to the beef industry.

1.6. THESIS HYPOTHESES

Transcriptomics, epigenomics, and metabolomics can reveal the impact of prenatal nutrition and/or RFI on complex biological interactions in young Angus bulls. The specific hypotheses for this study are as follows:

- The bovine metabolome can be comprehensively described and annotated using comprehensive, quantitative metabolomics techniques and computer-aided literature mining.
- Differences between HRFI and LRFI bulls can be detected using metabolomics techniques, such that serum-based metabolite biomarker panels can be developed to distinguish HRFI from LRFI animals.

- Differences in selection for parental RFI and/or prenatal nutrition can change gene expression patterns in tissues of young Angus bulls leading to changes in immunity and muscle development.
- Selection for parental RFI and/or differential prenatal nutrition can be associated with DNA methylation alterations in the sperm of young Angus bulls.

1.7. THESIS OBJECTIVES

My thesis has two aims: 1) understand some of the metabolic processes that underlie genetic differences in RFI, and phenotypic responses to prenatal nutrition, and their interactions in beef cattle; and 2) understand the influences of genetic differences in RFI and differences in prenatal nutrition on the sperm epigenome of young Angus bulls.

The first aim was divided into two phases, a metabolomic phase and a genomic phase. For the first phase, I used a variety of metabolomic techniques, including NMR spectroscopy, liquid chromatography-tandem mass spectrometry (LC-MS/MS), and ICP-MS to identify and quantify metabolites in four different bull tissues (*Longissimus thoracis* (LT) muscle, semimembranosus (SM) muscle, liver, testis) and two biofluids (serum and ruminal fluid). I hypothesized that differences in RFI and/or prenatal nutrition should be reflected in the metabolomic profile of biofluids and tissues in different animals. These differences should allow me to identify metabolic biomarkers for RFI prediction or potentially other metabolic biomarkers associated with other traits of interest such as immunity and muscle development. The use of multiple analytical platforms was intended to enhance the level of metabolome coverage with about 150 different metabolites (i.e. amino acids, fatty acids, carbohydrates, metals, etc.) being quantitatively measured in each tissue and biofluid sample. Metabolomic studies on these tissues and biofluids

were expected to reveal new insights into how RFI and/or prenatal diet could affect the metabolic profile of offspring.

For the second phase of aim 1, I used qPCR (real time-PCR or RT-PCR) to determine whether RFI and/or prenatal nutrition could affect the gene expression patterns of those imprinted and non-imprinted genes that have important roles in terms of immunity and myogenesis. For this phase of the study, the samples of interest were LT muscle, SM muscle, liver, and testis. This work was expected to lead to the identification of genetic biomarkers for RFI prediction or biomarkers associated with immune function as well as muscle development. Therefore, using metabolomics together with genomics it should be possible to reveal certain key aspects of the biology behind RFI influences and/or prenatal nutrition on subsequent bull calf traits. This work may also identify biomarkers associated with the traits of interest to select animals without phenotyping them. If these gene/metabolite biomarkers could be validated in a larger animal population, they could potentially be used in breeding programs to select for feed-efficient animals.

The second project aim used whole genome bisulfite sequencing (WGBS) on bull sperm to identify differentially methylated regions (DMRs) between HRFI and LRFI as well as between Normal-diet (Ndiet) and restricted or Low-diet (Ldiet) bulls. Sperm WGBS may also indicate how RFI or prenatal nutritional effects can pass on to the next generation. The identified DMRs were further interrogated to identify the biological pathways affected by parental RFI or the prenatal diet. Understanding the biological pathways that could be impacted by prenatal diet or RFI may help to develop nutritional or pharmaceutical manipulations that could improve prenatal dietary requirements or feed efficiency in beef cattle.

1.8. THESIS OUTLINE

This thesis consists of seven chapters. The current chapter provides a detailed literature review of RFI and maternal nutrition and their effects on progeny, livestock production and related livestock health. It also provides background material on the various "omics" technologies used to investigate the effects of RFI and prenatal maternal nutrition in many animal studies. Chapter 1 also provides the background, motivation, hypotheses and aims for this thesis.

The second chapter provides the results of the comprehensive, quantitative metabolomic characterization of six bovine biofluids and tissues, including serum, ruminal fluid, liver, LT muscle, SM muscle, and testis, using NMR spectroscopy, LC-MS/MS, and ICP-MS platforms. It also provides a detailed literature review of all the metabolites identified in different bovine tissues and biofluids. All of this information, plus detailed metabolite descriptions, referential NMR and MS spectral data and richly annotated biochemical pathway data has been placed in a publicly available database called the Bovine Metabolome Database (http://www.bovinedb.ca/). This chapter was published in the journal "Metabolites".

The third chapter compares the metabolic profile of bovine serum in LRFI vs. HRFI bulls. It describes the use of NMR spectroscopy, and LC-MS/MS techniques to characterize the metabolic differences between LRFI and HRFI animals. It also identifies potential serum biomarkers for predicting residual feed intake in young Angus bulls.

The fourth chapter investigates the impact of prenatal nutrition and selection for parental RFI on selected gene expression in young Angus bulls. In particular, in order to explore how RFI and/or prenatal nutrition impact the expression of genes associated with muscle development and immunity in the progeny Angus bulls, 26 imprinted and non-imprinted genes associated with muscle development and/or inflammatory response were selected from the literature. Expression

of these genes in LT and SM muscles, liver, and testis were tested by qPCR. This resulted in the identification of potential genetic biomarkers of feed efficiency and prenatal nutrition in cattle. This chapter was submitted to the journal "Livestock Science" for publication.

The fifth chapter investigates the impact of prenatal nutrition and selection for residual feed intake on sperm DNA methylation pattern in young Angus bulls. Using whole genome bisulfite sequencing (WGBS), DMRs between HRFI and LRFI as well as between Ndiet and Ldiet bulls were identified in the sperm of young Angus bulls. The identified DMRs were further interrogated to identify the biological pathways associated with them with respect to parental RFI and the prenatal diet treatments.

The final chapter consists of a short conclusion that summarizes the results of these studies, their interpretation and proposed future studies, which may reveal further insights into novel RFI biomarkers, improved assessments of maternal nutrition, the effects of these factors on progeny and their influence on beef cattle production and efficiency.

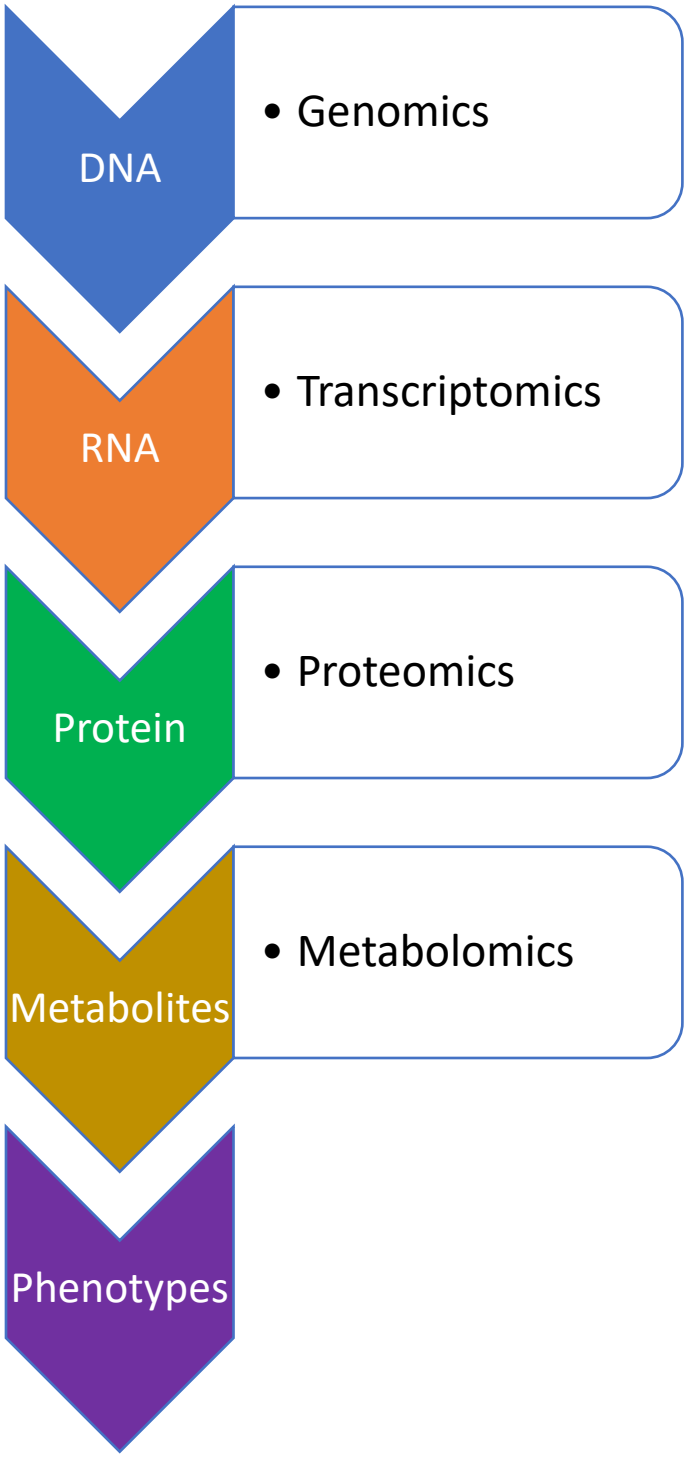

Figure 1. 1 General schema showing the relationships of the genomics, transcriptomics, proteomics, and metabolomics.

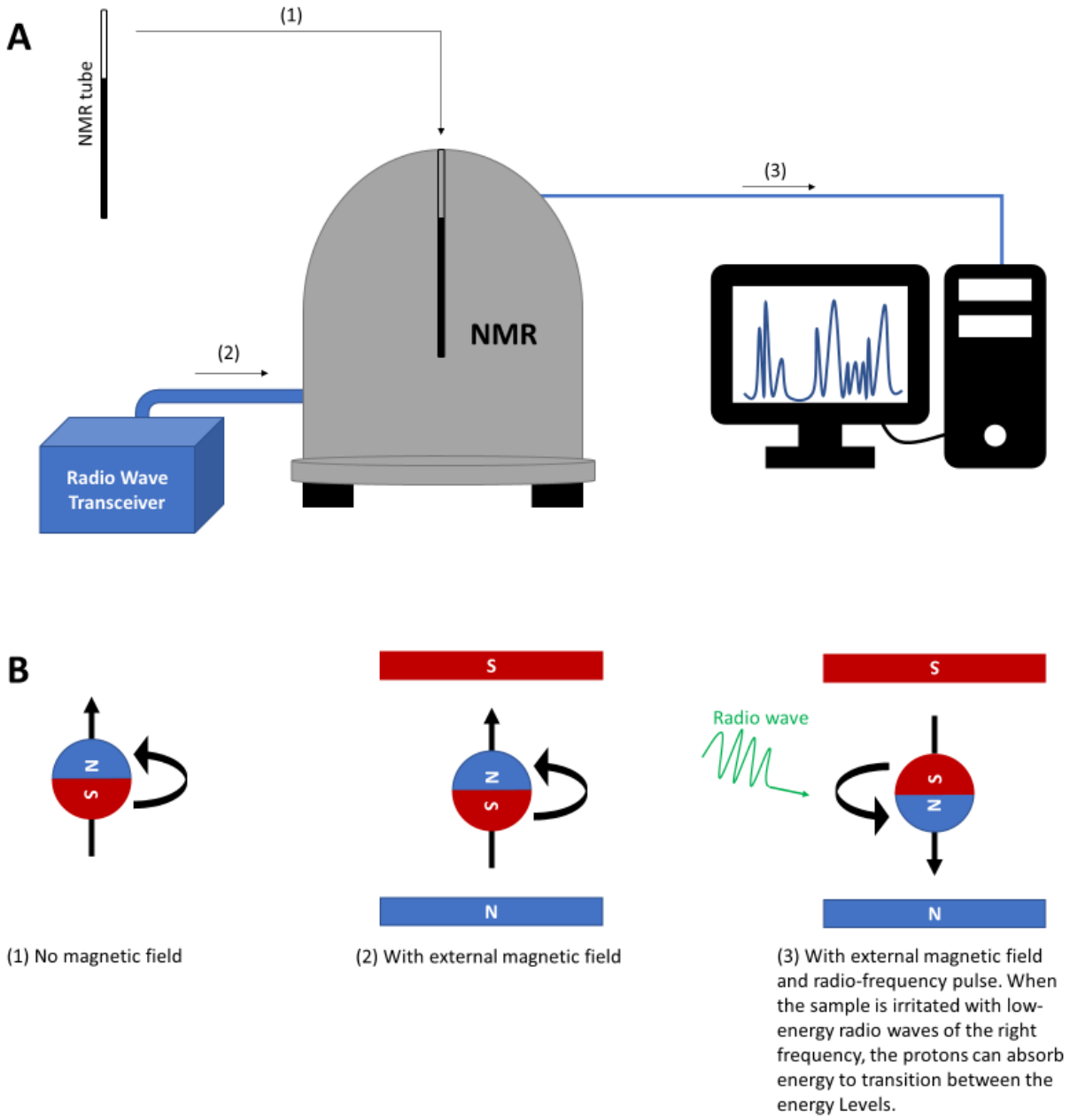

Figure 1. 2 A) Schematic diagram of how an NMR spectrometer works. 1) The NMR tube containing the biofluid or extracted tissue is placed into a magnetic field; 2) a weak oscillating magnetic field is passed through the sample; 3) This radio-frequency magnetic oscillation leads certain excited nuclei within the molecule to resonate and absorb the magnetic energy and to emit a detectable electromagnetic signal with a frequency characteristic of the magnetic field at the nucleus. B) Schematic diagram of how a magnetic field orients the dipoles in a magnet and how dipoles flip due to the radio-frequency pulses.

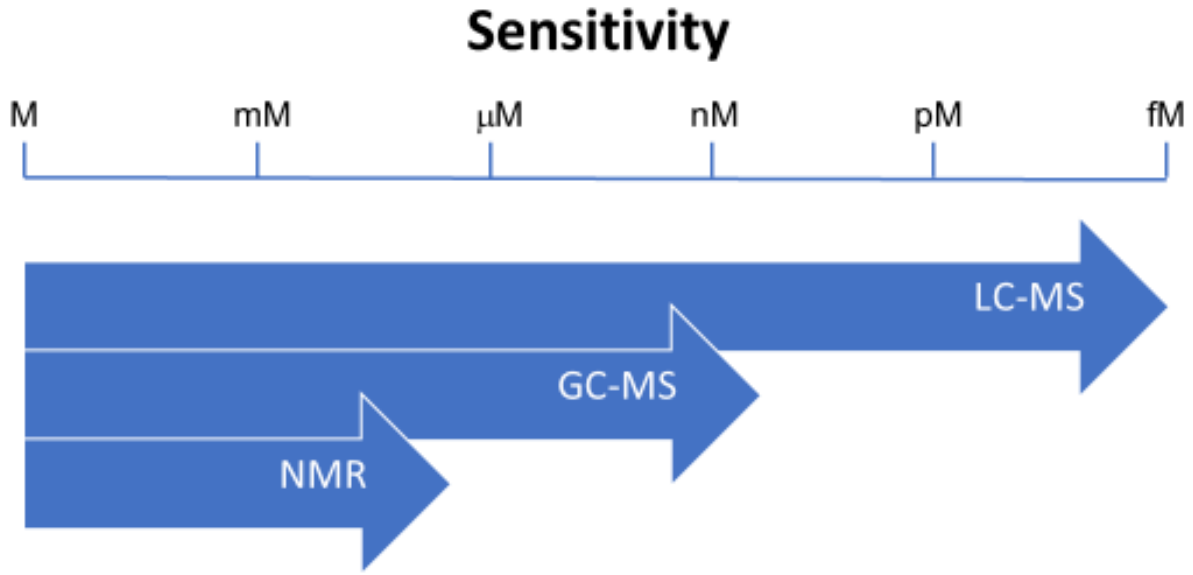

Figure 1. 3 Relative sensitivity of different analytical methods (NMR, GC-MS and LC-MS) used in metabolomics. NMR is the least sensitive method (limit of detection ~ 5 µM) and LC-MS is the most sensitive (limit of detection <1 nM).

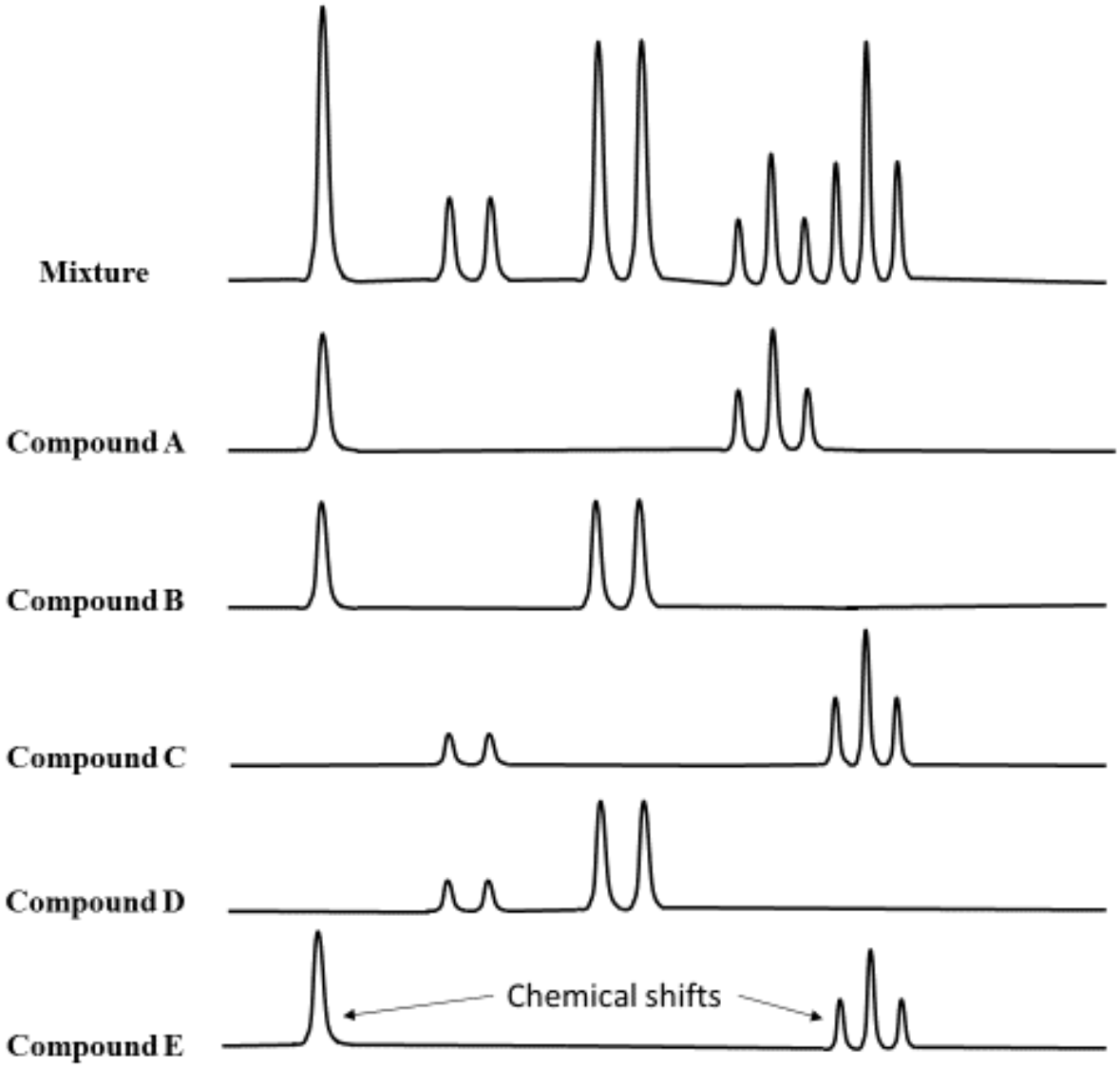

Figure 1. 4 The principles behind spectral fitting in NMR. The upper spectrum is a mixture composed of five compounds (A, B, C, D, and E), each with a unique spectrum containing different chemical shifts at different locations.

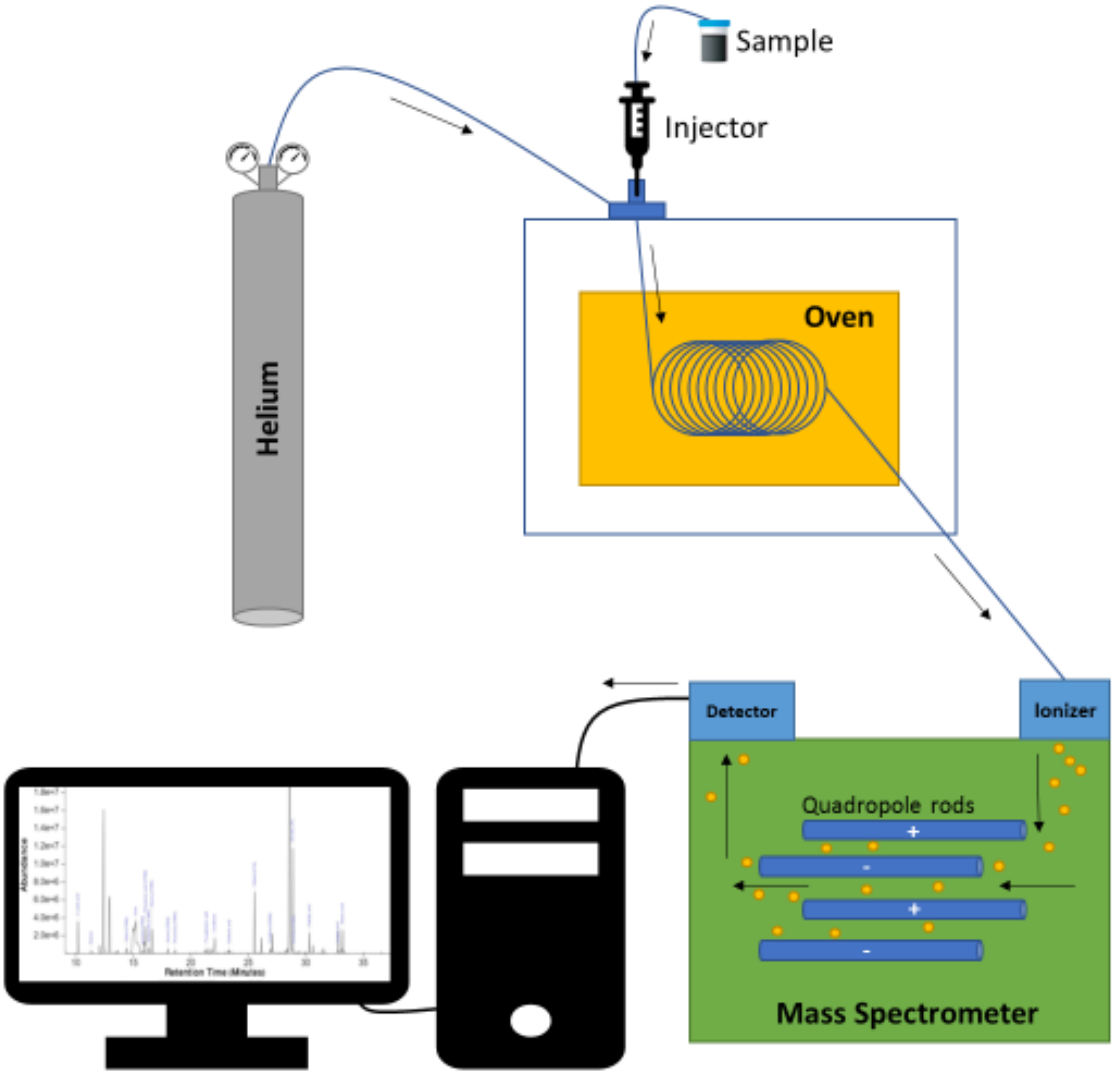

Figure 1. 5 Schematic view of how a GC-MS instrument works. A more detailed description of how GC-MS works is given in the text.

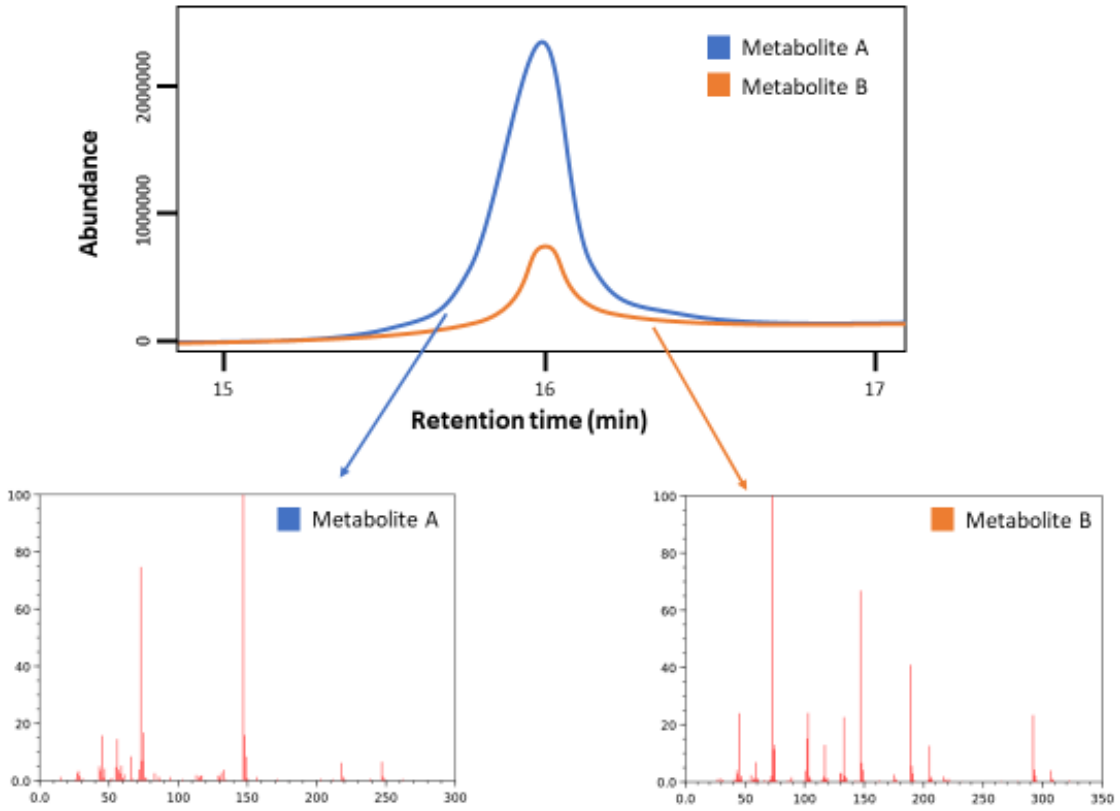

Figure 1. 6 Overview of how spectral deconvolution works in GC-MS. A GC-MS spectrum or total ion chromatogram for a biological sample typically consists of dozens to hundreds of sharp peaks (corresponding to ion counts) covering an elution time of approximately 30-45 min. Each peak may consist of one or more electron ionization (EI) mass spectra arising from one or more compounds. Similar principles apply in LC-MS.

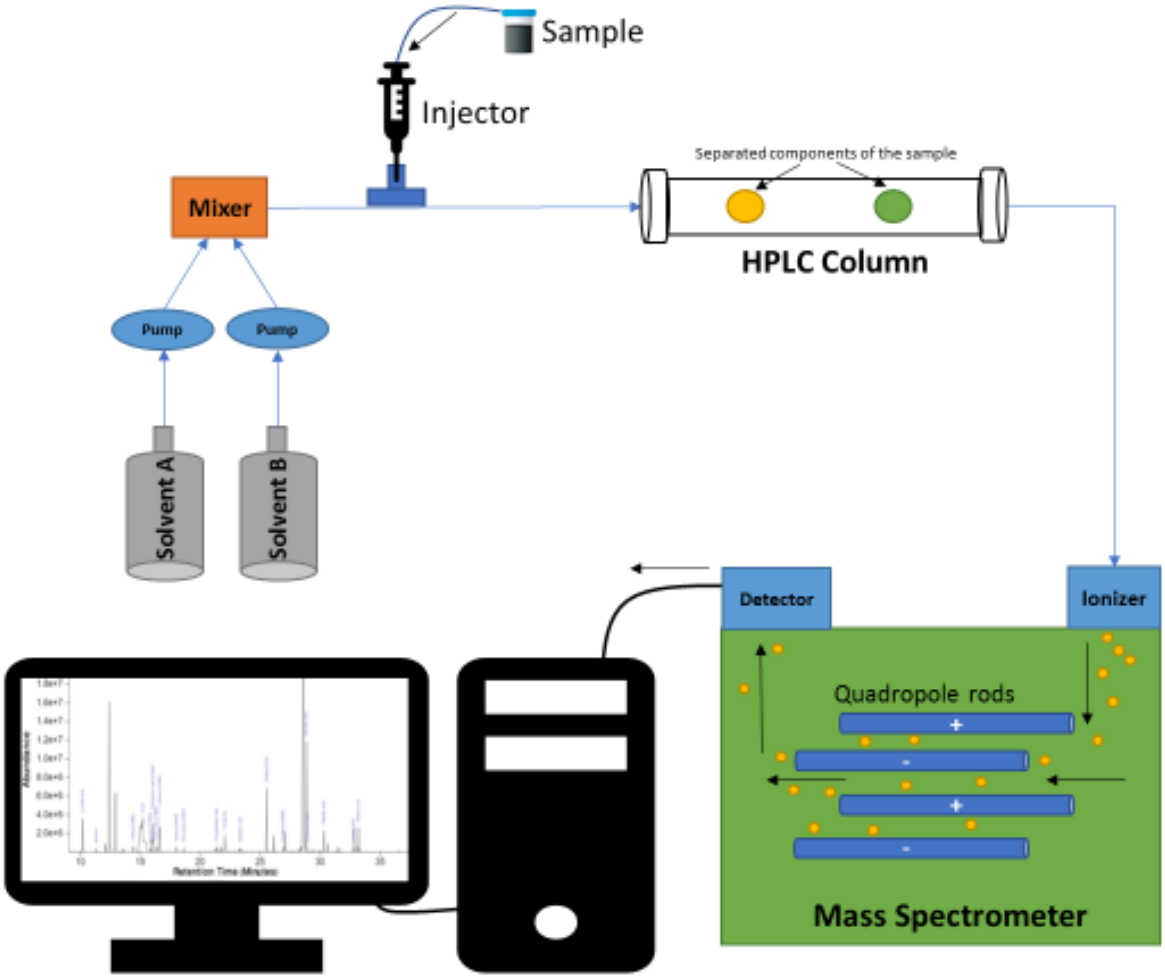

Figure 1. 7 Schematic view of how an LC-MS instrument works. A more detailed description of how LC-MS works is given in the text.

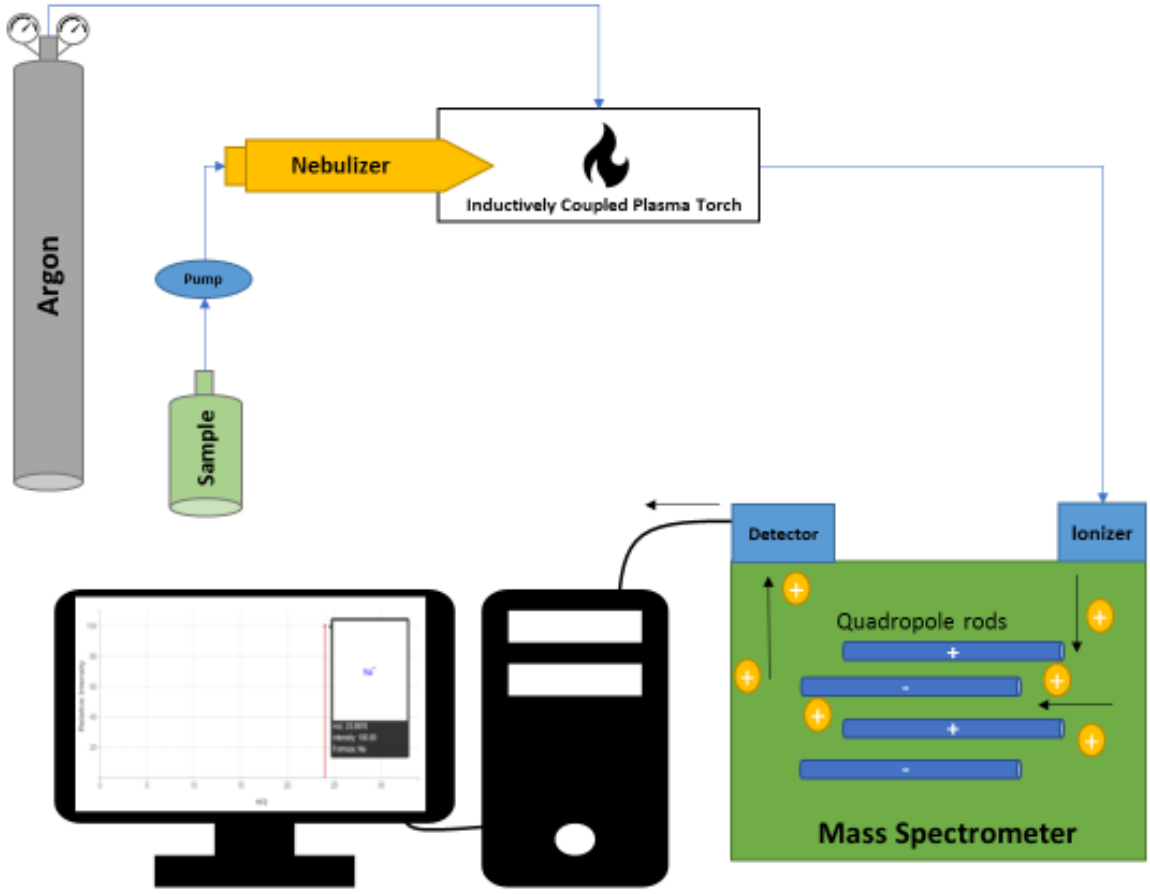

Figure 1. 8 Schematic view of how an ICP-MS instrument works. A more detailed description of how ICP-MS works is given in the text.

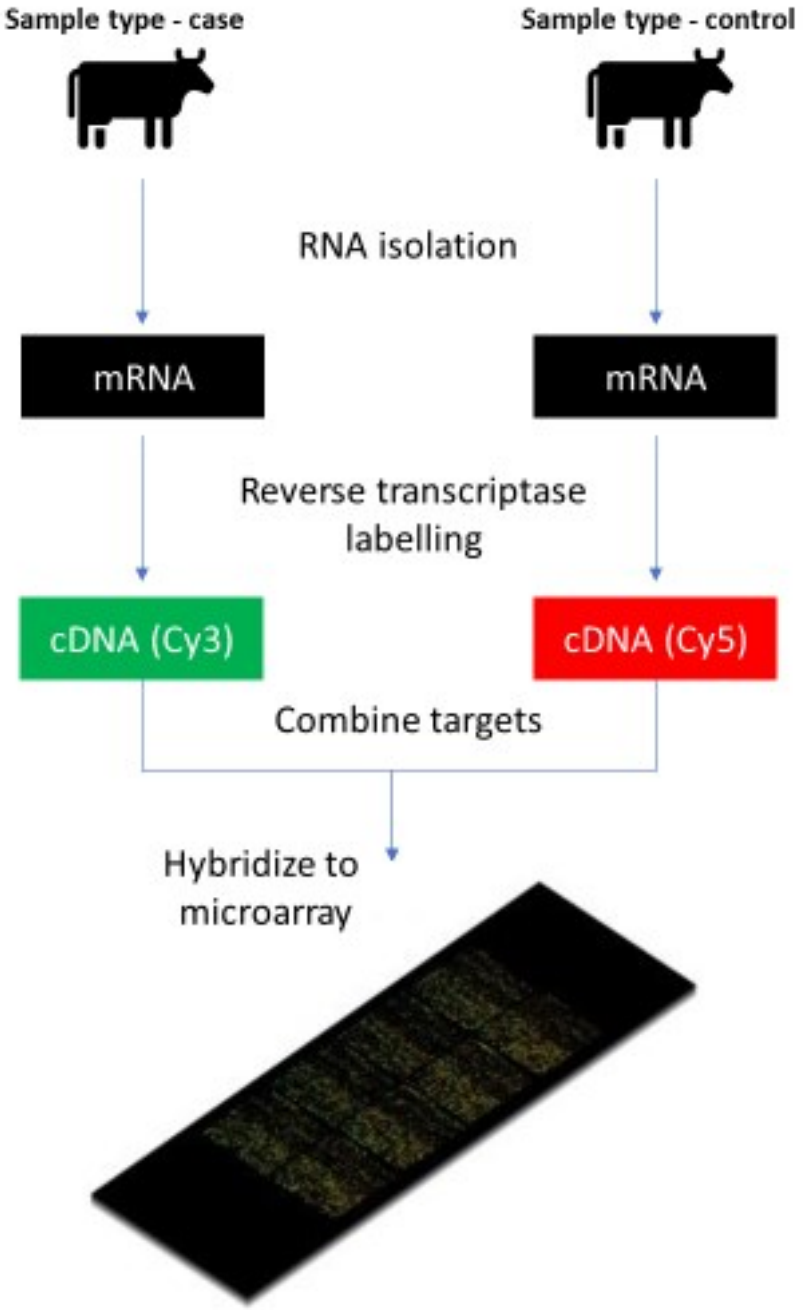

Figure 1. 9 Overview of the working principles of a typical two-color microarray experiment. RNA is isolated from a biological sample, the isolated RNA is labelled with Cy3 and Cy5 fluorescence dyes, and then the labeled mRNA is hybridized with cDNA probes in the microarray slide.

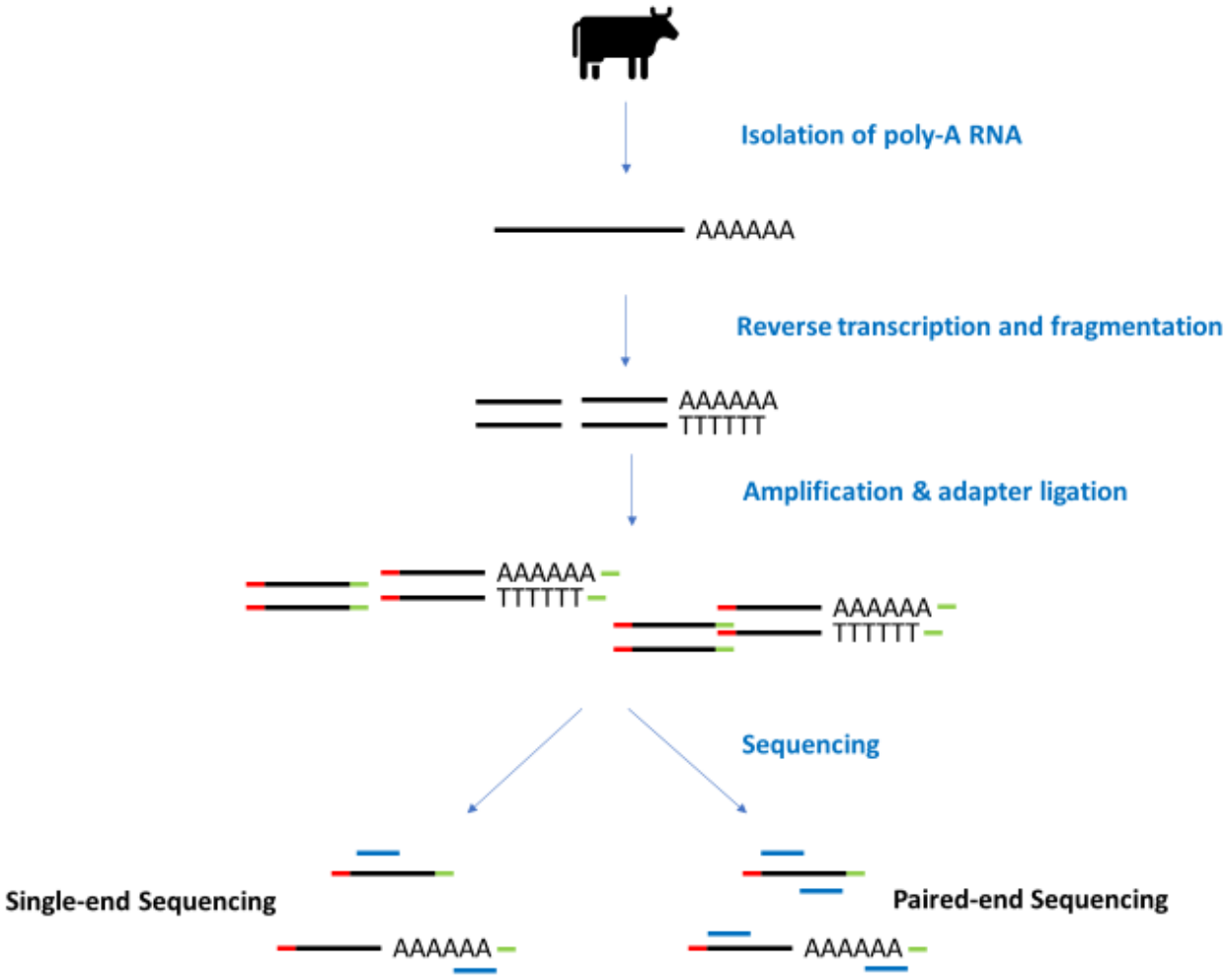

Figure 1. 10 Overview of a typical RNA-Seq workflow. RNA is isolated from a biological sample, the isolated RNA is reverse-transcribed, fragmented and amplified, the amplified fragments are ligated to adapters and then sequenced either single- or paired-end.

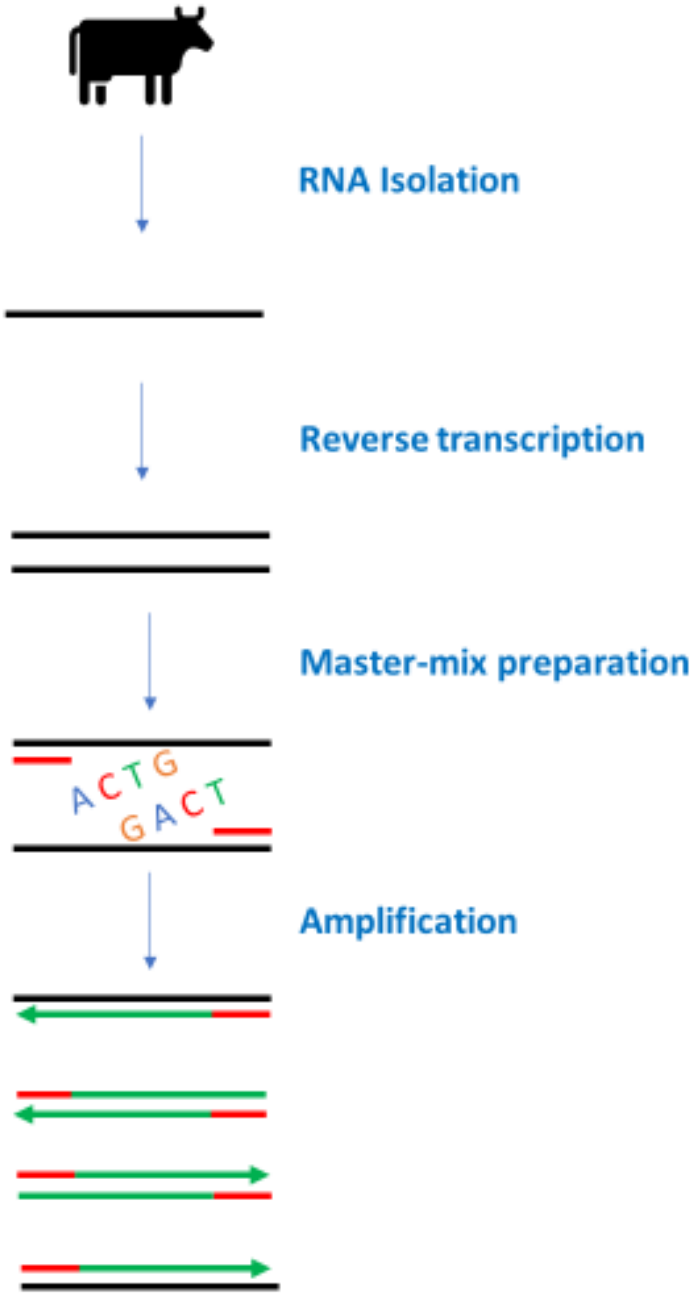

Figure 1. 11 Overview of a typical qPCR workflow. RNA is isolated from a biological sample, the isolated RNA is reverse-transcribed to make cDNA, the synthesized cDNA is mixed with PCR master-mix and then amplified using a qPCR machine.

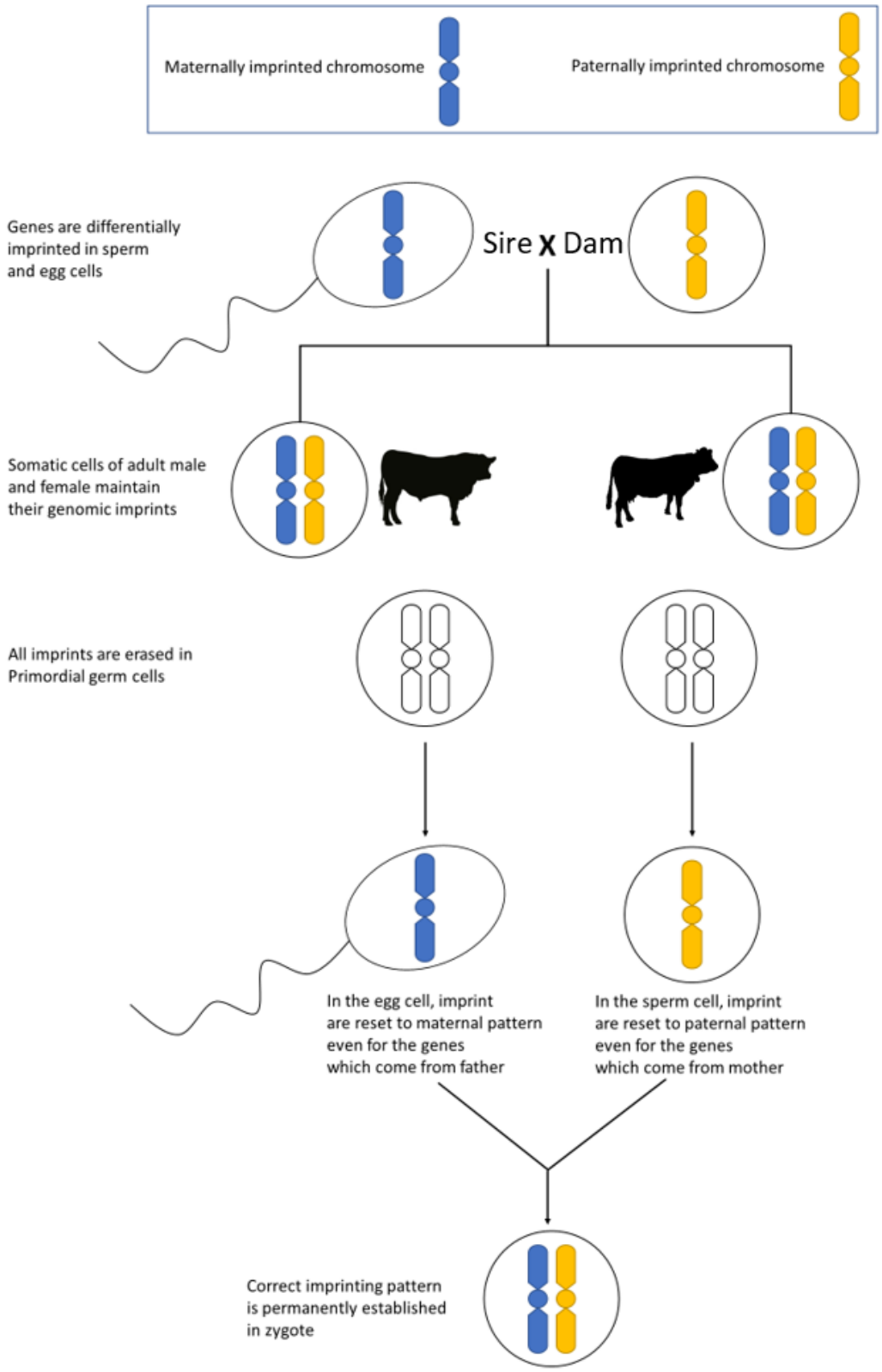

Figure 1. 12 Overview of how imprinting marks are transferred from parents to the next generation.

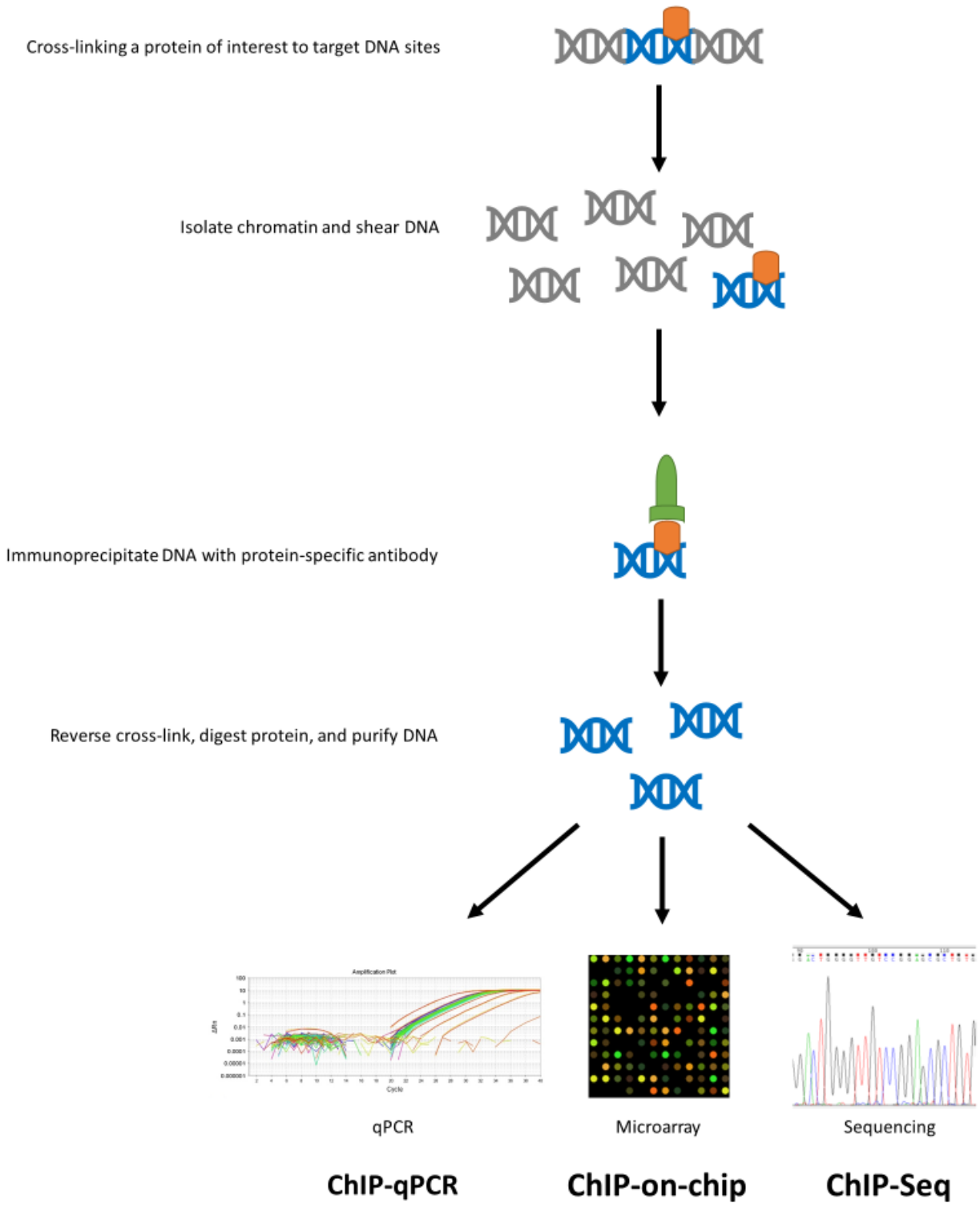

Figure 1. 13 Overview of a typical chromatin immunoprecipitation (ChIP) workflow coupled with qPCR (ChIP-qPCR), microarray (ChIP-on-chip), and sequencing (ChIP-Seq).

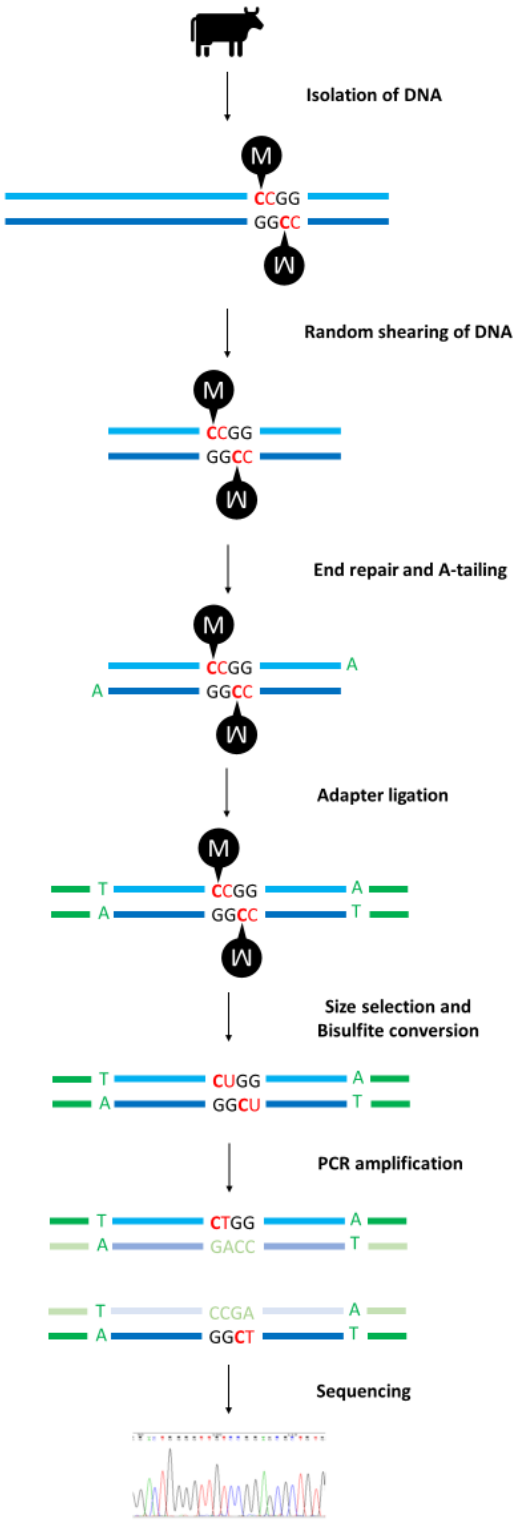

Figure 1. 14 Overview of a typical whole genome bisulfite sequencing (WGBS) workflow. DNA is isolated from a biological sample, the isolated DNA is fragmented, end-repaired, A-tailed, ligated with adapters, bisulfite treated, and amplified by PCR to make a library, then the library is sequenced on an NGS platform.

CHAPTER 2. THE BOVINE METABOLOME

2.1. INTRODUCTION

The cattle industry is among the most significant agri-food sectors in the world. It has been estimated that the global beef industry is worth more than CAD 300 billion/year [1] and is responsible for producing and processing > 70 million tonnes/year of meat [2]. The global dairy industry is worth > CAD 650 billion/year [3] and produces more than 800 million tonnes/year of milk or milk products [4], with more than 80% of those products coming from dairy cows [5]. Beef or beef products as well as milk and milk products are rich and dense sources of vital nutrients. They are nutritionally important food staples for hundreds of millions of people around the world. Indeed, the widespread use of bovine milk and bovine meat has been responsible for significant improvements to human health, prosperity and longevity over the past 200 years [6–11]. While the macronutrient (protein, fat, etc.) content of beef and milk is well known and has been studied for many decades, somewhat less is known about the micronutrient and chemical composition of key bovine biofluids and tissues [8,12,13]. Furthermore, from an animal health perspective, even less is known about the typical or healthy ranges of concentrations for many clinically important metabolites in bovine biofluids and tissues [14–17]. Indeed, far less is known about the metabolome of dairy and beef cattle than the metabolomes of other organisms, such as humans [18–21], yeast [22,23], bacteria [24,25] or even common crop plants, such as rice or tomatoes [26–28].

Over the past ten years, our knowledge gaps regarding the bovine metabolome have been slowly filled in through a variety of metabolomics studies focused on characterizing the chemical composition of different bovine biofluids and tissues. In particular, our laboratory has contributed significantly to the current state of knowledge of the bovine metabolome. Our first study, published

in 2013, focused on the characterization of the bovine ruminal metabolome [12]. This work, which used a combination of nuclear magnetic resonance (NMR) spectroscopy, inductively coupled plasma–mass spectrometry (ICP–MS), gas chromatography–mass spectrometry (GC–MS), and reverse-phase liquid chromatography–mass spectrometry (RPLC–MS) coupled with direct flow injection (DFI)–mass spectrometry (DFI–MS) (RPLC–DFI–MS/MS), led to the characterization and quantification of 246 ruminal fluid metabolites or metabolite species, including amino acids, biogenic amines, carbohydrates, lipids, vitamins, and trace minerals. Here, we define "metabolite species" as those molecules with non-unique chemical formulas or masses (such as lipid isomers-PC(36:6)) while "unique compound structures" correspond to compounds with a unique and clearly defined chemical structure and a unique chemical name (such as L-alanine). Subsequent studies done by our lab have explored the metabolite composition of bovine serum, ruminal fluid, and urine, leading to the identification of 142 metabolites in serum [14–17], 232 metabolites in ruminal fluid [12,29], and 52 metabolites in urine [17] using NMR, GC–MS, liquid chromatography–mass spectrometry (LC–MS) and/or ICP–MS. Most recently, we completed the most extensive metabolomics study ever done on bovine milk [13]. This study used a combination of NMR spectroscopy, LC–MS, and ICP–MS to identify and quantify 296 bovine milk metabolites or metabolite species (corresponding to 1447 unique structures). A further literature review identified 676 milk metabolites or metabolite species (corresponding to 908 unique structures), bringing the total to 972 metabolites or metabolite species [13]. Many other bovine metabolome studies have also been undertaken and published by other laboratories around the world [7,30–35]. To date there have been 45 metabolomics studies performed on bovine milk, 42 metabolomics studies on bovine serum/plasma, 25 metabolomics studies performed on bovine ruminal fluid, 4

metabolomics studies on bovine urine and more than 20 studies performed on a variety of bovine tissues, secreta and biofluids including liver, muscle, testes, and others [8,12,13,29,36,37].

While the number of metabolomics studies on bovine fluids and tissues has grown significantly, the majority of the studies published to date have focused on identifying rather than quantifying metabolites. Furthermore, most of these bovine-oriented metabolomics studies do not use more than two analytical techniques nor do they attempt to integrate previously published metabolite information to extend or validate their results. An additional challenge facing bovine metabolomics researchers lies in the fact that most of the metabolomics data is not consolidated or centralized into readily accessible public resource. To facilitate further research into the metabolome of beef and dairy cattle, we believe it is critical to create an open, publicly accessible resource that contains current, comprehensive and quantitative data on the bovine metabolome—including metabolomics data on multiple biofluids, excreta and tissues. We also believe that this resource should include detailed compound descriptions, information on referential LC–MS, GC–MS and NMR spectra, detailed biochemical pathway data and other data typically found in referential metabolome databases (refer to the Human Metabolome Database (HMDB) [18–21], the *Escherichia coli* Metabolome Database (ECMDB) [24,25], and the Yeast Metabolome Database (YMDB) [22,23]). Such an undertaking would benefit beef and dairy researchers, food scientists, nutritionists and consumers.

To create such a resource, we decided to combine experimental metabolomics techniques with computer-aided text mining to compile essentially all of the known chemical compounds (endogenous and exogenous) that can be detected in bovine milk, blood, urine, ruminal fluid, muscle, liver, and testes (as well as other biofluids and tissues) along with their respective concentrations. The resulting database is called the Bovine Metabolome Database (BMDB). The

BMDB, which is housed at www.bovinedb.ca [38], is a comprehensive web-accessible database containing concentration data, physico-chemical data and reference data for 51,801 bovine metabolites.

This paper describes the experimental, computational and literature research efforts used to collect and validate the metabolomics data in the BMDB as well as the techniques used to construct the BMDB and place it online. Experimentally, we focused on updating or expanding the data for several incompletely characterized biofluids or tissues, including bovine serum, ruminal fluid, liver, muscle, and testes. To do so, we used multiple quantitative metabolomics techniques, including high-resolution NMR spectroscopy, liquid chromatography–tandem mass spectrometry (LC–MS/MS) and ICP–MS methods. This experimental work allowed us to greatly add to the previously collected metabolomics data for these biofluids and tissues. To further enhance our experimental metabolic profiling studies, we conducted an extensive literature survey and extracted additional metabolite data from more than 240 journal articles identified through computer-aided literature searches spanning not only these biofluids/tissues but another six bovine biofluids and tissues. This "bibliomic" effort yielded data for another 1086 metabolites or metabolite species (which corresponds to 163 lipids and 923 non-lipids). The experimentally acquired metabolite data was then combined with genome-scale metabolite inference—a technique commonly used to fill in the metabolic "holes" for other metabolomes, such as the human metabolome [18–21], the yeast metabolome [22,23], and the *E. coli* metabolome [24,25]. This method uses known, organism-specific metabolic pathways [39] and known, organism-specific gene/protein reactions to provide data on metabolites that are known to exist, but not normally measured via NMR or MS techniques. This led to the addition of another 48,628 metabolites or metabolite species. All of these data (metabolite names, structures, descriptions, concentration

data, biofluid/tissue locations, physico-chemical data, and known or predicted NMR and MS spectra) were then placed in the BMDB. The BMDB is a web-accessible, MySQL-based database constructed using a Ruby-on-Rails web framework. The BMDB offers a variety of user-friendly data search, browse and display options similar to other popular metabolomics databases, such as the HMDB [18–21], YMDB [22,23], and ECMDB [24,25].

Overall, the intent of this study is to help address four key questions: 1) What kinds of compounds and nutrients are present in various bovine biofluids and tissues? 2) What is the approximate variation in the concentration of metabolites across different kinds of bovine biofluids and tissues? 3) What fraction of the bovine biofluids and tissues metabolome can be identified and/or quantified using targeted, quantitative metabolomics techniques? 4) What analytical methods (NMR, LC–MS/MS, ICP–MS) are best suited for comprehensively profiling the bovine metabolome? Answering these questions will provide a common foundation and a more appropriate set of reference values for both ongoing and future bovine biofluids and tissues composition studies.

2.2. MATERIALS AND METHODS

This study consisted of both "wet" (experimental) and "dry" (computational/literature) research. The "wet" component focused on the comprehensive, quantitative metabolomics characterization of six bovine biofluids and tissues using multiple metabolomics platforms. The "dry" component consisted of computer-aided literature research to complement and complete the collection of metabolomics data on the eight selected biofluids/tissues as well as four other bovine biofluids and tissues. It also included computational, genome-scale inference of biochemically expected bovine metabolites. The dry component also consisted of designing, constructing and testing the electronic

BDMB database. This section will describe the materials and methods for both the wet and dry components of the study.

2.2.1. Ethics approvals

The collection and analysis of bovine tissues and biofluids in this study were approved by the University of Alberta's Animal Care Committee (Animal Use Protocol (AUP) 1129) under the auspices of the Canadian Council of Animal Care [67].

2.2.2. Animal selection

Twenty-six purebred Angus bulls raised on the Roy Berg, University of Alberta Kinsella Ranch (Kinsella, AB, Canada), were used in this study. After weaning, the bulls were fed and managed according to industry standards for feedlot production of finished cattle in Alberta until they reached approximately 17 months of age [68]. The bulls were then slaughtered at the Agriculture and Agri-Food Canada (AAFC)-Lacombe Research Centre abattoir over a period of four days to collect the target tissues [68].

2.2.3. Sample collection

Two bovine biofluids including serum and ruminal fluid, and four bovine tissues including liver, LT muscle, SM muscle, and testis tissues were collected and experimentally characterized for this study. Serum and ruminal fluids were collected from live animals, whereas the tissues were collected from recently slaughtered animals. Blood samples (10 mL) were collected in the morning (just before feeding) at 15 months of age from a jugular vein using vacutainer serum collection tubes (Becton Dickinson, Mississauga, ON, Canada). Blood samples were kept in a cooler on ice, transferred to the laboratory within 3 hours after collection, and centrifuged at 2000 ×g at 4 °C for 15 min. The upper layer of serum was then collected, and 4 mL was stored at −80 °C. Ruminal

fluid samples were collected at 9 months of age through a rumen oro-ruminal tube in the morning prior to feeding. A 15 mL sample of ruminal fluid for each animal was obtained and placed on ice, then centrifuged at 3000 ×g for 20 min at 4°C to remove particular matter. The supernatant was then transferred into a 5 mL tube and centrifuged at 9400 ×g for 20 min at 4 °C for further phase separation. The clear upper phase from each tube was transferred into two 2 mL Eppendorf tubes and subsequently stored at −80 °C. To collect tissue samples (muscle, liver, and testes), the animals were slaughtered at age 17 months. Between 5 and 10 g tissue samples from LT muscle (from the left side of the animal between the 12th and 13th ribs), SM muscle (left), liver, and testes were collected approximately 30–45 min after death. The tissue was immediately frozen in liquid nitrogen and then stored at −80 °C.

2.2.4. Biofluid sample preparation for NMR

Biofluid sample preparation was done according to the procedure described by Foroutan et al. [69]. Both serum and ruminal fluid samples contain a substantial portion of large molecular weight proteins and lipoproteins which can seriously compromise the quality of ^{1}H-NMR spectra though the generation of intense, broad lines that interfere with the identification and quantification of lower abundance metabolites. De-proteinization can eliminate these peaks. The de-proteinization of these samples was done by ultrafiltration using 3-kDa ultrafiltration units (Amicon Micoron YM-3; Sigma-Aldrich, St. Louis, MO, USA), following a previously reported de-proteinization procedure [70]. Briefly, a newly opened 3 kDa Amicon 0.5 mL ultrafilter system was thoroughly rinsed by adding 500 µL HPLC-grade water (Millipore Sigma, Oakville, ON, Canada) to the filter and centrifuging at 9400 ×g for 10 min. The glycerol-containing filtrate was disposed into the sink followed by repeating the rinsing step four more times. The rinsed Amicon ultrafilter was then dried with a Kimwipe tissue and placed into a new 1.5 mL microcentrifuge tube. The biofluid

sample (either ruminal fluid or serum), which was thawed on ice (for 30 minutes), was briefly centrifuged at 9400 ×g for 3 minutes at 4 °C to remove any particulate material. The supernatant (450 µL) was then transferred into the Amicon ultrafilter and centrifuged at 11,500 ×g for 20 min at 4 °C to remove the proteins from the sample. The de-proteinized sample was then frozen and stored at −80 °C until further use.

2.2.5. Tissue sample preparation for NMR

For each tissue, approximately 2 g of a frozen sample was ground in a 300 mL pestle and mortar with about 100 mL liquid nitrogen. A total of 300 mg of the ground (still frozen) tissue was weighed and transferred into an 11 mL PYREX™ screw cap glass tube (Corning Inc., New York, NY, USA), followed by adding 4.4 mL cold methanol and 0.68 mL cold HPLC grade water. The tissue was homogenized using a Vortex-Genie 2 vortexer (Scientific Industries, New York, NY, USA) at 1500 rpm for 3 min, then 2.2 mL cold chloroform was added to the sample. The mixture was vortexed at 1500 rpm speed for 5 min and then centrifuged for 10 min at 1000 ×g at 4 °C. The supernatant was then transferred into a new 11 mL PYREX™ glass tube, then 2.2 mL cold chloroform and 3.2 mL cold water were added. The mixture was vortexed at 1500 rpm speed for 3 min, then centrifuged for 10 min at 1000 ×g at 4 °C. This will give a biphasic mixture. The upper polar phase (containing water-soluble metabolites) and lower non-polar phase (containing lipid-soluble metabolites) were carefully separated using a Pasteur pipette. The lower non-polar phase was transferred into a 2 mL PYREX™ glass tube and kept at −80 °C for future metabolomics analysis. The upper polar phase was transferred into a 15 mL Falcon tube (Thermo Fisher Scientific, Whitby, ON, Canada). The upper polar phase was purged under nitrogen gas for 90–120 min. Once the purging was completed, 2 mL of HPLC-grade water was added to the tube. The tube was then frozen with liquid nitrogen and the sample lyophilized for 24 h. After lyophilization,

the dried sample was dissolved with 300 µL of HPLC-grade water and kept at −80 °C until further use.

2.2.6. NMR spectroscopy

Two-hundred-eighty µL of the ultrafiltered biofluid (serum or ruminal fluid) or the water-soluble extract of each tissue was transferred to a 1.5 mL Eppendorf tube, to which an additional 70 µL of a standard NMR buffer solution was added. For serum, the buffer consisted of 250 mM potassium phosphate (pH 7.0), 5 mM 2,2-dimethyl-2-silapentane-5 sulfonate (DSS-d_6), 5.84 mM 2-chloropyrimidine-5-carboxylic acid, and D_2O 54% v/v in H_2O. For ruminal fluid, liver, SM and LT muscle and testis tissue extracts, the buffer consisted of 750 mM potassium phosphate (pH 7.0), 5 mM 2,2-dimethyl-2-silapentane-5 sulfonate (DSS-d_6), 5.84 mM 2-chloropyrimidine-5-carboxylic acid, and D_2O 54% v/v in H_2O. The mixture (a final volume of 350 µL) was then transferred to a 3 mm NMR tube for spectral analysis. All ^{1}H-NMR spectra were collected on a Bruker Avance III Ascend 700 MHz spectrometer equipped with a 5 mm cryo-probe (Bruker Biospin, Rheinstetten, Germany). ^{1}H-NMR spectra were collected at 25 °C using the first transient of a nuclear Overhauser effect spectroscopy (NOESY)-presaturation pulse sequence. This pulse sequence was selected based on its excellent reproducibility and quantitative accuracy [71]. NMR spectra were acquired with 128 scans employing a 4 second acquisition time and a 1 second recycle delay.

2.2.7. NMR compound identification and quantification

Prior to spectral deconvolution, all free induction decays (FIDs) were zero-filled to 240,000 data points and a 0.5 Hz line broadening function was applied. The methyl singlet of the added DSS (set to 0.00 ppm) served both as an internal chemical shift referencing standard and as an internal

standard for quantification. All ^{1}H-NMR spectra were processed using the Chenomx NMR Suite 8.1 software package (Chenomx Inc., Edmonton, AB, Canada) for compound identification and quantification as previously described [72]. A minimum of three experienced NMR spectroscopists processed and analyzed each NMR spectrum to eliminate compound identification and quantification errors. Sample spike-in experiments were also used to confirm the identity of a number of compounds suspected to be present in specific biofluids or tissue samples. A spike-in experiment involves adding 50–500 μM of the presumptive compound to selected samples to test if the corresponding ^{1}H-NMR signals changed as expected. NMR analysis typically led to the identification and quantification of about 50 metabolites in each biofluid or tissue sample.

2.2.8. LC–MS/MS compound identification and quantification

A targeted, quantitative metabolite profiling approach was employed that combined RPLC–MS with DFI–MS (RPLC–DFI–MS/MS) to determine the concentrations of a wide range of metabolites. These analyses were performed using an in-house quantitative metabolomics assay (TMIC Prime) [73,74]. This assay was used with an Agilent 1260 series ultrahigh-performance liquid chromatography (UHPLC) system (Agilent Technologies, Palo Alto, CA) coupled with an AB SCIEX QTRAP® 4000 mass spectrometer (Sciex Canada, Concord, ON, Canada) to identify and quantify up to 143 compounds (including amino acids, biogenic amines, glucose, organic acids, acylcarnitines, PCs, LysoPCs, SMs, and SM(OH)s). The absolute quantification of water-soluble compounds including amino acids, organic acids, and biogenic amines was ensured by using two separate UHPLC injections with C18 column separations. On the other hand, glucose and the various lipid classes (acylcarnitines, PCs, LysoPCs, SMs, etc.) are measured by the column-free DFI method (both +ve and –ve mode). While initially designed and calibrated for human metabolomics studies, the measurable ranges of metabolite concentrations available

through the TMIC Prime assay match very closely with the known or expected metabolite concentrations in bovine biofluids and tissues (as determined via orthogonal NMR experiments and high levels of agreement with published literature data).

The detection of each metabolite in the TMIC Prime assay relies on multiple reaction monitoring (MRM). The assay incorporates both isotope-labelled internal standards and other quality control (QC) standards into its 96-well filter plate to ensure accurate compound quantification. The first 14 wells in the 96-well plate are used for building calibration curves and QCs, while the other 82 wells are used for sample analysis. For all biofluids analyzed with this assay, both the original sample (without dilution) and diluted samples (10×) were analyzed to ensure correct calibration and quantification. In brief, 10 µL of each sample (the filtered biofluid or the water-soluble or lipid-soluble extract of the tissue) was carefully pipetted into an appropriate sample well of the upper 96-well filter plate and dried using a stream of nitrogen gas. Amino acid and biogenic amine derivatization were done by adding 50 µL of a 5% solution of phenyl-isothiocyanate (PITC) to each well and incubating for 20 minutes. After incubation and PITC derivatization, the samples were dried down using a nitrogen gas evaporator. The metabolites were then extracted by adding an ammonium acetate/methanol mixture (5 mM ammonium acetate in methanol) to the upper 96-well filter plate, shaking at 330 rpm for 30 min, and then centrifuging the plates so that the extract bled into the lower 96-deep well plate.

The resulting extract was split for RPLC–MS (150 µL) and DFI–MS (150 µL) analyses followed by a dilution step with 150 µL of water for RPLC–MS analysis and with 400 µL of the MS running solvent for DFI–MS analysis. All LC–MS analyses were conducted on an AB SCIEX QTRAP® 4000 mass spectrometer equipped with an Agilent 1260 series UHPLC system. The Analyst software 1.6.2 (Concord, ON, Canada) was used to control the entire assay's workflow.

2.2.9. Trace elemental analyses using ICP–MS

All trace elemental analysis was performed on a Perkin-Elmer NexION 350x ICP–MS (Perkin-Elmer, Woodbridge, ON, Canada), operating in a kinetic energy discrimination (KED) mode. Argon (ICP/MS grade, 99.99 %) was used as a nebulizer (0.9 mL min^{-1}), an auxiliary (1 mL min^{-1}) and a plasma gas (15 mL min^{-1}). Helium (He) was used as a non-reactive collision gas (Cell gas A: 4.3) to eliminate/minimize chemical interference. The dwell time for each metal ion was set to 50 ms with a total integration time of 500 ms (10 sweeps per reading and three replicates). The uptake of samples/standards/QCs was done by a peristaltic pump using the following protocol: 1) sample flush for 50 seconds at 48 rpm, 2) read delay for 15 seconds at 20 rpm, 3) spectral analyses at 20 rpm, and 4) washing for 45 seconds at 24 rpm. All samples were diluted using 1% HNO_3, 5% H_2O_2, and MiliQ water (grade 1) by a factor of 10 for serum and tissues and a factor of 5 for ruminal fluid samples. Indium (In) was added to the dilution solvent as an internal standard. The final concentration of indium after mixing with the samples/standards was 20 ppb. An external calibration curve was used for the quantitation of all metal ions using a six- to nine-point calibration curve (for each metal) and linear regression. The performance of the ICP–MS was checked daily using a commercially prepared Perkin Elmer Nexion calibration solution to evaluate the sensitivity of the instrument. The Nexion solution was also used to calibrate the mass spectrometer at low (Be), mid (In), and high (U) masses. The accuracy of the ICP–MS analytical protocol was evaluated in every sequence by the analysis of standard reference materials (SRMs) - i.e., serum and water QCs. Continuing calibration verification (CCV) was run every 15 samples to monitor the validity of each calibration curve throughout the sequence.

2.2.10. Literature research on bovine biofluid and tissue metabolites

We conducted an extensive literature review of known bovine biofluids and tissues metabolites and their concentrations using many open-access search engines, such as Google Scholar [75], PubMed [76], and ScienceDirect [77]. We also used several in-house text-mining software packages that were originally developed for the Human Metabolome Project (HMP) and the Human Metabolome Database (HMDB) [78]. Two of the most useful programs were PolySearch [79] and PolySearch2 [80]. These programs are able to take simple keywords (i.e., "serum", "bovine", etc.) as input and rapidly create hyperlinked lists of abstracts and papers from PubMed (and other data sources) containing information about bovine metabolites and their corresponding concentration data. PolySearch2 was able to compile a ranked list of bovine metabolites by measuring word co-occurrence frequency using terms such as "cow serum", "serum", "beef", "bovine", and "cattle" in conjunction with words such as "concentration", "identification", "quantification", "mM", or "micromol". PolySearch2 also extracted key sentences from the abstracts, then labelled and hyperlinked the metabolites mentioned in the text. This led to the identification of ~200 papers, abstracts and textbooks with relevant chemical information on bovine biofluids and tissues.

All literature-derived data regarding bovine compounds, along with their concentrations and references, were compiled, compared, and their names "normalized" to match HMDB [78], chemical abstracts service (CAS) number, and PubChem identifiers. The manually derived compound data was further annotated using an in-house program called DataWrangler [78], which automatically generates names, synonyms, descriptions, structures, chemical taxonomies, physical property data, and bioavailability data. The information generated by DataWrangler was manually checked by three different scientists with post-graduate degrees in biochemistry, physiology,

and/or animal sciences. Additional data compiled from two other electronic databases, including the Livestock Metabolome Database (LMDB) [36] and the Milk Composition Database (MCDB) [13] was also collected and cross-checked. After the manual checking phase was complete, the data were then entered into the BMDB. Concentration data were cross-checked manually to identify any large discrepancies (>3×) between entered values. Those that exceeded this threshold were re-analyzed to see if data entry errors had been made. For highly discrepant values, a "majority wins" scheme was used to select the best or most likely value. On the other hand, if our experimental data matched best with one of the discrepant values, that value was selected over other reported value(s). The resulting list of 14 literature-derived bovine biofluids and tissues metabolites (including 278 overlapping experimentally-derived metabolites), along with their concentration data (when available), helped to confirm many of the metabolites and metabolite concentrations previously found in our experimental analyses.

2.2.11. Genome scale inference of expected bovine metabolites

Publicly available bovine metabolite, protein and pathway data from PathBank [39] and UniProt [81] were used to generate a genome-scale compilation of biochemically inferred or "expected" bovine metabolites. These "expected" metabolites correspond to endogenously produced compounds that are very likely to be in cells or tissues, based on well-known or well-characterized biochemical pathways or reactions. Many of these compounds correspond to lipids, transient intermediates, or low-abundance compounds that are not normally measured in metabolomics experiments. Comparisons of these computationally inferred bovine metabolites with metabolite and protein data reported in the HMDB was used to identify potentially missing metabolites or metabolite species. Care was taken to exclude exogenous compounds (human-only food additives, drugs, food constituents, etc.) that would be unlikely to be found in cattle or cattle feed.

2.2.12. Construction of the BMDB

BMDB was implemented using a Ruby on Rails (http://rubyonrails.org, version 4.2.0) web framework incorporating a MySQL relational database (https://www.mysql.com, version 15.1 Distrib 10.4.6-MariaDB) to manage all of the metabolite data, including descriptions, synonyms, physico-chemical properties, concentrations, spectra, and external references. BMDB was built using HMDB's framework and is therefore similar in appearance and structure. BMDB uses the model–view–controller architecture, in which internal data logic is separated from user input and data presentation. The raw information stored in the database is dynamically extracted and rendered into web pages. BMDB is hosted on a Digital Ocean server equipped with 4 CPUs, 80 GB of disk space, and 8 GB of RAM.

2.3. RESULTS

This section is divided into four subsections covering: 1) experimental metabolomics results; 2) literature review results; 3) a comparative assessment between different tissues, biofluids, and platforms; and 4) a detailed description of the BMDB.

2.3.1. Water-soluble compound identification and quantification by NMR and LC–MS/MS

Using a combination of NMR and LC–MS/MS, a total of 58, 64, 60, 60, 66, and 69 water-soluble compounds were identified and quantified in serum, ruminal fluid, *Longissimus thoracis* (LT) muscle, semimembranosus (SM) muscle, liver, and testis tissues, respectively. The complete list of compound concentrations (including averages and their standard deviations) for serum is shown in Table 2.1. The most abundant compounds in serum were lactate (average concentration: 4.8 ± 2 mM), glucose (4.0 ± 0.4 mM), and urea (1.3 ± 0.3 mM). The lowest concentration that could be reliably detected in serum was 0.04 ± 0.02 μM for putrescine. The complete list of compound

concentrations for ruminal fluid (including averages and their standard deviations) is shown in Table 2.S1. The most abundant compounds in ruminal fluid were acetate (37 ± 8 mM), butyrate (26 ± 9 mM), propionate (16 ± 5 mM), and glucose (16 ± 11 mM). The lowest concentration in ruminal fluid that could be reliably detected was 0.1 ± 0.1 µM for serotonin.

In terms of LT and SM muscle tissues, the complete list of compound concentrations (including averages and their standard deviations) is shown in Table 2.S2 The most abundant compounds detected in these muscle tissues, in terms of µmol per gram wet-weight were lactate (LT = 31 ± 8 µmol/g, SM = 38 ± 11 µmol/g), carnosine (LT = 22 ± 5 µmol/g, SM= 22 ± 3 µmol/g), creatine (LT = 4.7 ± 0.4 µmol/g, SM= 4.8 ± 0.3 µmol/g), and glutamine (LT = 2.8 ± 0.8 µmol/g, SM = 2.5 ± 0.7 µmol/g). The lowest concentration that could be reliably detected was for spermidine (0.09 ± 0.03 nmol/g in LT muscle and 0.11 ± 0.04 nmol/g in SM muscle) as well as spermine (0.2 ± 0.1 nmol/g in LT muscle and 0.11 ± 0.03 nmol/g in SM muscle).

Two other non-muscle tissues were analyzed, including liver and testes. The complete list of compound concentrations (including averages and their standard deviations) for liver tissue is shown in Table 2.S3. The most abundant compounds in liver, in terms of µmol per gram wet-weight, were glucose (80 ± 15 µmol/g), lactate (12 ± 1.7 µmol/g), and sn-glycero-3-phosphocholine (11 ± 1.5 µmol/g), while the least abundant compound was putrescine (0.2 ± 0.1 nmol/g). The complete list of compound concentrations (including averages and their standard deviations) for testis tissue is shown in Table 2.S4. The most abundant compounds in testis, in terms of µmol per gram wet-weight were lactate (7.7 ± 1.6 µmol/g), creatine (7.6 ± 1.9 µmol/g), myo-inositol (7.2 ± 1.2 µmol/g), O-phosphoethanolamine (7 ± 1.2 µmol/g), and glutamate (3.3 ± 0.7 µmol/g), whereas the least abundant compound was spermidine (0.4 ± 0.2 nmol/g).

2.3.2. Lipid-like compound identification and quantification by LC–MS/MS

A locally developed LC–MS/MS assay called The Metabolomics Innovation Centre (TMIC) Prime assay provided quantitative results for 74 lipid-like compounds (14 lysophosphatidylcholines (LysoPCs), 5 sphingomyelins (SMs), 5 hydroxysphingomyelins (SM(OH)s, 10 phosphatidylcholines (PCs), and 40 acylcarnitines) for serum (Table 2.1), LT and SM muscles (Table 2.S2), liver (Table 2.S3), and testis (Table 2.S4). Sixty-four of these lipid-like compounds (10 LysoPCs, 3 SMs, 3 SM(OH)s, 8 PCs, and 40 acylcarnitines) were detected and quantified in ruminal fluid (Table 2.S1). The other 10 compounds that could be quantified in serum and tissue samples, but not in ruminal fluid, were below the limit of detection (LOD) in ruminal fluid samples. Note that each LysoPC and PC species identified by the TMIC Prime assay typically corresponds to a minimum of two up to 24 possible unique structures, respectively. In our study, SM(16:0) (68 ± 10 μM) and LysoPC(14:0) (6 ± 3 μM) were the most abundant lipids identified in serum and ruminal fluid, respectively. Carnitine or C0 (1856 ± 424 nmol/g and 1751 ± 337 nmol/g) was the most abundant metabolite identified in LT and SM muscle, respectively. Likewise, carnitine (22 ± 6 nmol/g) and SM(16:0) (22 ± 7 nmol/g) were the most abundant lipid-like compounds identified by the TMIC Prime assay in liver, whereas SM(16:0) (42 ± 6 nmol/g) was the most abundant lipid-like compound in testis. The least abundant lipids detected by TMIC Prime assay were acylcarnitines. These included C16:2-OH (5 ± 1 nM) and C18:2 (6 ± 1 nM) in serum and ruminal fluid, respectively. They also included a single acylcarnitine, C18:2, at 0.004 ± 0.001 nmol/g, 0.004 ± 0.001 nmol/g, 0.006 ± 0.001 nmol/g, and 0.004 ± 0.001 nmol/g for LT muscle, SM muscle, liver, and testis, respectively.

2.3.3. Trace element identification and quantification by ICP–MS

The ICP–MS method was developed to detect and quantify 35 metabolites or trace minerals. While no toxic trace metals (i.e., Pb, As, Cd) were detected in the samples, ICP–MS provided quantitative results for 13 metabolites or trace minerals in bovine serum (Table 2.1). The most abundant trace elements identified and quantified by ICP–MS were sodium (134 ± 14 mM), potassium (4 ± 0.4 mM), calcium (2 ± 0.2 mM), and phosphorus (1.3 ± 0.2 mM), while the least abundant metals quantified by ICP–MS were cesium (1.7 ± 0.3 nM), barium (200 ± 30 nM), and strontium (1000 ± 100 nM). In terms of ruminal fluid, our ICP–MS analysis provided quantitative results for 17 metabolites or trace minerals, as shown in Table 2.S1. The most abundant trace elements identified and quantified were sodium (236 ± 20 mM), potassium (40 ± 7 mM), phosphorus (12 ± 2 mM), and magnesium (7.5 ± 3.5 mM), whereas the least abundant compounds were cesium (30 ± 10 nM) and cobalt (1000 ± 200 nM).

ICP–MS could identify and quantify 15, 14, and 15 trace minerals in bovine LT muscle, SM muscle, and liver, respectively (Table 2.S2 and Table 2.S3). In LT and SM muscles, the most abundant trace elements identified and quantified were potassium (LT = 49 ± 8 μmol/g; SM = 52 ± 7 μmol/g), phosphorus (LT = 22 ± 3 μmol/g; SM = 22 ± 3 μmol/g), and sodium (LT = 10 ± 2 μmol/g; SM = 12 ± 3 μmol/g), whereas the least abundant compounds were thallium (LT= 0.0007 ± 0.0002 nmol/g; SM= 0.0009 ± 0.0001 nmol/g) and vanadium (LT = 0.013 ± 0.004 nmol/g; SM= 0.009 ± 0.001 nmol/g). In liver, the most abundant trace elements identified and quantified were potassium (46 ± 8 μmol/g), phosphorus (33 ± 7 μmol/g), and sodium (26 ± 5 μmol/g), while the least abundant compounds were cesium (0.02 ± 0.01 nmol/g), and lead (0.019 ± 0.003 nmol/g). For testis tissue, ICP–MS could identify and quantify 16 metabolites or trace minerals (Table

2.S4). The most abundant trace elements identified and quantified were potassium (49 ± 8 μmol/g), sodium (36 ± 7 μmol/g), and phosphorus (16 ± 3 μmol/g). The least abundant compounds were thallium (0.0014 ± 0.0003 nmol/g) and lead (0.014 ± 0.003 nmol/g).

2.3.4. The chemical composition of bovine biofluids and tissues (experimental data)

Inspection of our experimental data reveals that the chemical composition of bovine serum is dominated by inorganic ions (primarily sodium, potassium, calcium, and phosphorus), carbohydrates (glucose), organic acids (lactate, acetate, and 3-hydroxybutyrate), amino acids (glycine, valine, and glutamine), and various amine-containing compounds (urea, creatinine). Ruminal fluid mostly contains inorganic ions (primarily sodium, potassium, phosphorus, and magnesium), carbohydrates (glucose), organic acids (acetate, butyrate, and propionate), amino acids (lysine, glutamate, and alanine), biogenic amines (methylamine and putrescine), and nucleobases (uracil). The LT and SM muscles are mostly composed of inorganic ions (primarily potassium, phosphorus, and sodium), carbohydrates (glucose), organic acids (lactate), amino acids (creatine, glutamine, and alanine), biogenic amines (carnosine) and miscellaneous compounds (O-acetylcarnitine and betaine). The most abundant compounds in liver are inorganic ions (primarily potassium, phosphorus, and sodium), carbohydrates (glucose), organic acids (lactate), amino acids (glutamate and glycine), biogenic amines (taurine) and miscellaneous compounds (sn-Glycero-3-phosphocholine). The most abundant chemicals in testis are inorganic ions (primarily potassium, sodium, and phosphorus), carbohydrates (UDP-N-acetylglucosamine), organic acids (lactate), amino acids (creatine and glutamate), biogenic amines (O-phosphoethanolamine and taurine), and miscellaneous polyols (myo-inositol). Lesser quantities of lipids including acylcarnitines (except carnitine and acetylcarnitine, which have high concentrations in both LT and SM muscles),

LysoPCs, PCs, SMs, as well as other small bioactive compounds are also evident in these bovine biofluids and tissues.

2.3.5. Literature survey of bovine biofluids and tissues metabolites

As part of our literature survey, 249 papers were reviewed on bovine metabolomics. These papers provided metabolomic data for six biofluids and eight tissues. The paper that provided the most extensive metabolomics data was written by Foroutan et al. [13] on the milk metabolome. This paper identified and quantified 972 metabolites (296 metabolites were experimentally detected and 676 were found from the literature) found in bovine milk using a combination of LC–MS/MS, NMR, ICP–MS and literature reviews [13]. Other papers of note include the work of Saleem et al. [12], which described more than 200 metabolites found in bovine ruminal fluid (using a combination of NMR, GC–MS, ICP–MS, and LC–MS/MS), the paper by Zhang et al. [16], which identified and quantified 128 metabolites found in serum (using a combination of direct injection and tandem mass spectrometry (DI-MS/MS) with a reverse-phase LC–MS/MS), and the paper by Muroya et al. [8], which identified and quantified 70 metabolites found in LT muscle of Japanese Black (Wagyu) cattle (using Capillary Electrophoresis–Time-of-flight–Mass Spectrometry, also known as CE–TOF–MS). The most thoroughly studied bovine biofluid was milk with 118 papers, followed by blood with 35 papers [36]. The least studied bovine biofluid was semen with three papers [54–56]. The most thoroughly studied bovine tissue was muscle with 12 papers [7,8,36,57,58], whereas the least studied bovine tissue was liver, with just three papers [59–61]. The most comprehensively characterized bovine biofluid was milk with 972 metabolites identified via literature data [13], followed by ruminal fluid with ~ 200 metabolites [12].

As we observed with both our experimental data and from the literature data, bovine serum is particularly rich in inorganic ions or minerals such as sodium (107–137 mM), potassium (4–4.3

mM), calcium (1.4–2.2 mM), phosphorus (1.3–1.6 mM), magnesium (0.8–0.9 mM), as well as a variety of organic acids such as lactate (0.6–1.6 mM), acetate (0.9–1 mM), and 3-hydroxybutyrate (0.2–2 mM). Other highly abundant metabolites reported in bovine serum include glucose (3–4 mM), urea (2–4 mM), and acetone (0.1–1 mM). The least abundant compounds in serum include several acylcarnitines, such as C18:1-OH (8–9 nM), C16-OH (3–6 nM), as well as C3-OH and C7-DC and C14 (10–12 nM). The concentrations reported in the literature data for the above-mentioned metabolites agreed well with our experimental data, except for lactate (2.8–6.8 mM), acetate (0.2–0.6 mM), and urea (1.1–1.6 mM).

Just as with serum, bovine ruminal fluid is rich in a variety of inorganic ions or minerals, such as sodium (110–117 mM), potassium (17.9–18.2 mM), phosphorus (9.1–9.3 mM), calcium (0.9–1 mM), and magnesium (0.09–0.1 mM). The most abundant organic compounds reported in ruminal fluid are organic acids such as acetate (41–81 mM), propionate (14–17 mM), and butyrate (6–18 mM), amino acids such as proline (240–1275 μM), isoleucine (123–1210 μM), and lysine (91–1095 μM), as well as glucose (393–3111 μM). The least abundant compounds in ruminal fluid include several lipids or acylcarnitines such as C16:2 (2–13 nM), C18:1-OH (2–22 nM), C14:2-OH (8–32 nM), and C16-OH (9–11 nM), sphingomyelins such as SM(16:0) (10–50 nM), hydroxysphingomyelins such as SM (16:1(OH)) (10–20 nM) and SM (14:1(OH)) (10–30 nM), as well as the lysophosphatidylcholine, LysoPC16:0 (10–200 nM). The concentrations reported in the literature for the above-mentioned metabolites agreed well with our experimental data, except for sodium (216–255 mM), potassium (33.3–46.5 mM), phosphorus (10.5–14.3 mM), magnesium (3.9–10.9 mM), glucose (4746–27122 μM), SM(16:0) (130–1370 nM), SM(16:1(OH)) (30–110 nM), and SM(14:1(OH)) (40–60 nM). These differences are most likely due to dietary differences

among the different cattle groups being studied, as the concentration of many ruminal fluid metabolites are strongly affected by the feed given to cattle.

In terms of bovine LT and SM muscles, the literature data indicated that organic acids such as lactate (14–132 μmol/g), inosinate (0.031–9 μmol/g), and succinate (0.9–2.3 μmol/g), amino acids such as creatine (1–10 μmol/g), glutamine (3–5 μmol/g), and alanine (1–1.3 μmol/g), carbohydrates such as glucose (3.3 μmol/g) and glucose-1-phosphate (0.1–1.2 μmol/g), the biogenic amine, carnosine (10–14 μmol/g), as well as betaine (1.5–1.8 μmol/g) and carnitine (0.8–2.6 μmol/g) were found to be quite abundant in these tissues. The least abundant compounds in LT and SM muscles are several amino acids such as aspartate (14–54 nmol/g) and tryptophan (35–95 nmol/g), as well as the biogenic amine, putrescine (9–22 nmol/g). The concentrations reported in the literature for the above-mentioned metabolites agreed with our experimental data, except for glucose (LT: 0.3–0.8 μmol/g, SM: 0.3–1.2 μmol/g) and carnosine (LT: 17–27 μmol/g, SM: 19–25 μmol/g).

Regarding bovine liver tissue, a single study described by Miles et al. [59] was the only study that attempted to identify and quantify metabolites in the bovine liver tissue. The most abundant compound found was glutamate (6122–7999 nmol/g), followed by alanine (2366–3515 nmol/g) and glutamine (1911–2576 nmol/g). The least abundant compounds found in liver in this study was ornithine (984–1184 nmol/g). These values agreed moderately well with our findings.

We found literature values for 12 amino acids in bovine testis tissue. The most abundant metabolites reported in the literature for bovine testis are glutamate (2–2.4 μmol/g), glycine (0.9–1.3 μmol/g), and alanine (0.9–1.2 μmol/g), while the least abundant compounds are histidine (20 nmol/g), isoleucine (30 nmol/g), leucine (40 nmol/g), valine (50 nmol/g), and lysine (40–200 nmol/g). These values agreed moderately well with our findings with the observed differences

likely due to diet, age and breed, chemical volatility, sample work-up or extraction, sample storage protocols, and instrument sensitivity.

2.3.6. The BMDB website

All of the detected, identified, quantified and biochemically inferred compounds obtained from our experimental, computational, and literature-searching efforts have been deposited into a freely accessible database called the BMDB (Bovine Metabolome Database - www.bovinedb.ca). The database contains not only the metabolite names and synonyms (common and International Union of Pure and Applied Chemistry (IUPAC)), but also their structures (multiple formats), basic descriptions, chemical ontology, physico-chemical properties, their reference spectra (NMR, GC–MS, and LC–MS), pathway information (as derived from PathBank) [39], and literature citations from the scientific literature for all (to the best of our knowledge) of these compounds. The BMDB is structured to be very similar to other popular, online metabolomics databases, such as the HMDB [18–21], ECMDB [24,25] or YMDB [22,23]. In particular, the BMDB is designed to serve as a user-friendly tool for data browsing and compound searching. The BMDB can be launched from its home page (Figure 2.1) using a variety of pull-down menus or tabs located at the top of the home page, which serve as a navigation panel. Users may “Browse”, “Search”, or “Downloads” data from the database using this navigation panel. Additional information about the database is located under the “About” menu. Clicking on the “Browse” tab and the subsequent “Metabolites” dropdown option (on the BMDB navigation panel) generates a tabular view that allows users to casually scroll through the metabolites in the database or to re-sort its contents by compound name or mass. Each compound entry in the BMDB is hyperlinked to an individual metabolite description table (called a MetaboCard) that, when clicked, brings up additional information on that particular chemical. Clicking on the “Search” tab at the top of the home page allows users to perform

compound name searches, general text searches, compound structure searches, molecular weight searches, and a variety of spectral (NMR and MS) searches. The "Downloads" section contains several data files that include the chemical structures, metabolite data, spectra, and other files (in SDF, CSV, or XML format). Currently, the BMDB contains information on 3173 unique, experimentally detected compounds with unique, well-defined structures and names. The BMDB also contains another 48,628 compounds that have been computationally inferred to exist from detailed genomic analysis, biochemical pathway analysis, and comparison to other well-studied mammalian metabolomes. As a result, the BMDB contains three classes of metabolites: 1) detected and quantified; 2) detected but not quantified; and 3) expected (or genomically/biochemically inferred). These classes of metabolites may be easily selected or filtered using BMDB's filtering functions.

Note that the total number of chemical compounds in BMDB is not a number that will remain unchanged. Certainly, as technology and instrument sensitivity improve, it is anticipated that this number will increase as other, lower abundance metabolites will be detected and will be added to future versions of the BMDB.

2.4. DISCUSSION

Using a combination of experimental, computational and literature-based approaches, we have attempted to identify and quantify as many chemicals as possible that are detectable in different biofluids and tissues of both dairy and beef cattle. We have deposited this information into the BMDB (www.bovinedb.ca) [38]. The BMDB contains a total of 51,801 metabolites with unique compound structures. Just 4.1% of these metabolites with unique structures (corresponding to 11.8% metabolites or metabolite species) have concentration data in at least one tissue or biofluid, while the remaining 95.9% of these metabolites with unique structures (or 88.2% metabolites or

metabolite species) have no quantification information whatsoever. The metabolites in the BMDB have been associated with eight bovine tissues and six different bovine biofluids. Of the 2100 compounds with unique structures that were experimentally identified and quantified in the BMDB, 1834 (87.3%) were experimentally characterized by our laboratory (using NMR, GC–MS, LC–MS and ICP–MS methods) over the past 7 years, including 306 (14.6%), as reported in this communication. Another 266 (12.7%) metabolites were compiled from computer-aided literature searches and were therefore measured by other laboratories. In addition, 48,628 (93.9%) metabolites were biochemically inferred to exist through detailed comparisons to existing genomics/metabolomics data of other well-studied mammals.

There is considerable variability in the extent of metabolome coverage for different bovine tissues and biofluids. These results are summarized in Table 2.2. This table shows that the most fully characterized bovine biofluid is milk with 928 known metabolites or metabolite species (corresponding to 2350 unique structures). The most poorly characterized biofluid is urine with 62 known metabolites or metabolite species (corresponding to 62 unique structures). Likewise, the most completely characterized bovine tissue is liver with 1056 known metabolites or metabolite species (corresponding to 1254 unique structures) while the most poorly characterized tissues or biofluids are parathyroid gland, pineal gland, umbilical cord, and vitreous humor. These poorly characterized tissues are grouped in Table 2.2 under the "All tissues" category, each with just one known metabolite (corresponding to one unique structure).

Analysis of the data in the BMDB also allows us to identify which metabolomics platforms are most useful or most widely used for bovine metabolomics studies. According to our data, the most popular metabolomics platform for analyzing bovine metabolites is NMR, with 26 published studies. GC–MS is a distant second with 13 published studies, while LC–MS was used for 12

studies. ICP–MS has been used with just three studies. Other platforms used by metabolomics researchers include GC×GC–MS with 11 studies, and high-performance liquid chromatography–ultraviolet spectroscopy (HPLC–UV) with eight studies. In terms of metabolome coverage, LC–MS methods have been used to identify 472 compounds and to quantify 154 compounds. On the other hand, NMR methods have been used to identify 170 compounds and to quantify 108 compounds. GC–MS has been used to identify 425 compounds and quantify 41 compounds, while ICP–MS has been used to identify and quantify 44 compounds. Based on the list of chemicals identified by each platform there appears to be considerable overlap with the compounds identified by NMR and the compounds identified by GC–MS. On the other hand, LC–MS tends to permit the identification and quantification of more hydrophobic compounds. As mentioned earlier, the differences in metabolite coverage from the different platform technologies are largely due to differences in sensitivity, detection/instrument biases and separation protocols, as well as differences in compound stability, solubility, volatility and other intrinsic chemical factors [13].

2.4.1. Comparisons to other studies

One of the objectives of this work was to evaluate the level of agreement between different platforms and different laboratories (i.e., protocols) in identifying and quantifying key metabolites across identical biofluids and tissues and over different periods of time. Overall, we found a very good agreement between the results reported for most methods and most laboratories. Here, we define a very good agreement as being within one standard deviation of the reported literature value (i.e., literature value ± 1 SD). Moderate agreement corresponds to being within two standard deviations of the reported value, while a poor agreement is defined as a value that is greater than two standard deviations of the reported literature value. A small number of questionable or profoundly different, literature-derived concentration values were identified during our literature

survey, but most of these could be eliminated through the curation process after being deemed mistaken, disproven (by subsequently published studies), mistyped or physiologically impossible. Much of the curation process involved multiple curators carefully reading and double-checking the primary literature to annotate the concentration unit types, to perform unit conversions, and to catch typographical inconsistencies.

In terms of serum composition, we found that almost all studies exhibited very good agreement across platforms and across laboratories with almost no significant differences in either the composition or concentration of reported metabolites. More specifically, in terms of water-soluble compounds, an inspection of Table 2.1 reveals a generally good agreement between the NMR- and LC–MS-measured concentrations and those reported in the literature. Forty-one out of the 58 compounds identified by these two techniques in serum had concentration values previously reported in the literature. More than 73% (30/41) of these compounds exhibited good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value) in serum. ICP–MS is considered as a gold standard for the identification and quantification of trace metal ions [62], and, therefore, we tend to have higher confidence in the values derived via ICP–MS over those measured by other (older or less sensitive) technologies. In serum, all nine trace elements identified and quantified in our study exhibited very good agreement with previously reported literature values. In total, 64 out of the 145 compounds identified by NMR, LC–MS/MS, and ICP–MS had concentration values that were previously reported in the literature. More than 78% (50/64) of these compounds exhibited very good agreement with literature values for bovine serum. For instance, the value of alanine reported by our study ranged from 210–270 μM, and for the literature-derived data it ranged from 151–222 μM. This widespread agreement was not unexpected as serum/plasma must be highly stable and

cannot vary much in its metabolite concentrations to ensure physiological homeostasis [63]. Of course, there were a few exceptions to this rule. The most variable metabolite reported in serum was betaine. The value of betaine reported by our study ranged from 138–200 μM, and the literature-reported values ranged from 14–26 μM. This variation could be due to a number of factors, including differences in diet, sex, age, breed, sample work-up or extraction, sample storage protocols, analytical platforms and instrument sensitivity.

In contrast to serum and plasma, ruminal fluid exhibited considerable variability across platforms and laboratories, even after normalizing to creatinine. In terms of water-soluble compounds, the inspection of Table 2.S1 reveals a generally good agreement between the NMR- and LC–MS-measured concentrations and those reported in the literature. A total of 55 out of the 64 compounds identified in ruminal fluid had concentration values previously reported in the literature. More than 65% (36/55) of these compounds exhibited good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value) in ruminal fluid. Regarding trace elements, 15 out of the 17 the compounds identified in our study had concentration values previously reported in the literature. Five of the 17 compounds exhibited good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value). In total, 93 out of the 145 compounds identified by NMR, LC–MS/MS, and ICP–MS in the ruminal fluid had concentration values previously reported in the literature. More than 58% (54/93) of these compounds exhibited good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value) in ruminal fluid.

It was of some interest to compare our earlier ruminal fluid analysis done in 2013 [12] with the more recent analysis reported here. The most variable metabolites reported in ruminal fluid

were 3-phenylpropionate and magnesium. The value of 3-phenylpropionate reported by our 2020 study ranged from 0.03–0.07 mM, while the value reported in our 2013 study ranged from 0.3–0.7 mM. We re-analyzed our 2013 data and found a graphical error in fitting the 3-phenylpropionate's chemical shifts in NMR, which explains the concentration difference between these two studies. The value of magnesium reported by our 2020 study ranged from 4–11 mM, while the value reported in our 2013 study ranged from 0.09–0.10 μM. We speculate that the differences in the concentration of magnesium is mainly due to the dietary treatment differences in this study versus our earlier 2013 study.

In terms of tissues, we found generally good agreement between the results obtained for the metabolite composition of muscle, testis, and liver. For water-soluble compounds in muscle tissues, an inspection of Table 2.S2 reveals a generally good agreement between the NMR- and LC–MS-measured concentrations and those reported in the literature. A total of 43 out of the 60 compounds identified in both LT and SM muscles had concentration values previously reported in the literature. For the trace minerals identified in our study, we could not find any values reported in the literature. In total, 43 out of the 149 as well as 43 out of 148 compounds identified by NMR, LC–MS/MS, and ICP–MS in the LT and SM muscles, respectively, had concentration values previously reported in the literature. More than 65% (28/43) of these compounds exhibited very good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value) in muscle tissues. For instance, the value of beta-alanine reported for NMR ranged from 92–206 μM in LT muscle and 85–159 μM in SM muscle, whereas the literature-derived values ranged from 84–155 μM. The most variable compounds seen in muscle tissues were carnosine (LT: 17–27 μmol/g, SM: 19–25 μmol/g, as compared to Muroya et al. [8]: 10–14 μmol/g) and glucose (LT: 0.3–0.8 μmol/g, SM: 0.3–1.2 μmol/g, as compared to

Kim et al. [57]: 3.3 μmol/g). We suspect that these variations may be due to factors such as differences in diet, sex, age, breed, chemical volatility, sample work-up or extraction, sample storage protocols, and instrument sensitivity. Similarly, 12 out of the 159 compounds (12 out of 69 water-soluble compounds) as well as 4 out of 155 compounds (4 out of 66 water-soluble compounds) identified in the testis and liver tissues, respectively, had concentration values previously reported in the literature. Forty-one percent (5/12) of the compounds found in testis tissue exhibited very good agreement with literature values (i.e., the average values from our experiments fell within one standard deviation of the literature value). However, the magnitude of this variation was small in testis. As with other tissues, we suspect these small variations may be due to the same factors, such as differences in diet, sex, age, etc., mentioned above.

Interestingly, none of the four metabolites previously reported in liver (including alanine, glutamate, glutamine, and ornithine) exhibited very good agreement with the values we measured experimentally, although the magnitude of this variation was small. The reasons for this appear to be due to differences in sample processing and the fact that the values reported by Miles et al. [59] were measured via an HPLC coupled with a fluorometric detection instrument. Furthermore, a comparison of these values reported by Miles et al. [59] to metabolite values found in other tissues and other mammals [64–66] suggests that their values are likely incorrect.

2.4.2. Comparisons across platforms

We also found that, for those metabolites that were measured by both LC–MS/MS and NMR, there was a generally good overall agreement with the concentration values. Depending on the sample type, NMR and LC–MS/MS were able to identify a common set of 20–26 metabolites. Table 2.3 lists the concentrations measured by NMR and by LC–MS/MS for these common metabolites in the serum analyzed in this study. As can be seen from this table, the values differed by less than

10%, on average. The most significant exceptions were for sarcosine (28%) and glucose (14%) with the concentrations reported by NMR being 4 μM for sarcosine and 4.6 mM for glucose, whereas the concentrations reported by LC–MS/MS were 3 μM for sarcosine and 4 mM for glucose. Regarding sarcosine, the values are at the lower limit of sensitivity for NMR and so the differences are likely due to instrumental noise. However, glucose has 13 peaks (chemical shifts) in the NMR spectrum where some of these peaks overlap with other compounds (i.e., aspartate and taurine). As a result, these spectral overlaps can affect the concentration of glucose measured by NMR. Furthermore, these differences are likely a consequence of small differences in sample preparation or separation protocols, as well as differences in compound stability, solubility, volatility and other intrinsic chemical factors [13].

The metabolome coverage we achieved was maximized by using as many different platforms as possible and carefully optimizing the metabolite coverage of each platform. Some platforms clearly performed better than others. Depending on the type of sample, 39–49 compounds could be identified and quantified using NMR spectroscopy whereas 13–17 metal ions could be identified and quantified by ICP–MS. On the other hand, LC–MS/MS identified and quantified 102–116 compounds (depending on the sample type) of which 38–42 were non-lipid compounds and 64–74 were lipid or lipid-like compounds.

Because of their fundamentally different separation and detection technologies, different metabolomics platforms tend to target or detect different classes of metabolites. For instance, NMR is relatively “untargeted” but is biased towards highly abundant, water-soluble compounds. Other methods were quite targeted, with ICP–MS being limited only to metal ions and LC–MS/MS (the TMIC Prime assay) being limited to a pre-selected set of 143 compounds, including amino acids, biogenic amines, organic acids, and lipid-like compounds. While this study did employ a relatively

wide range of metabolomics platforms (NMR, LC–MS/MS, ICP–MS), it did not use all available detection tools (GC–MS, GC×GC-TOF), nor did it explore all available separation protocols (e.g., solid phase extraction and enrichment, immune or ELISA detection, chemical derivatization, etc.). However, for this particular study, we wanted to address the question of how well a cross section of commonly accessible metabolomics technologies or platforms could perform in identifying and quantifying metabolites in various bovine biofluids and tissues.

Overall, LC–MS/MS appears to be the most suitable method for the characterization of bovine biofluids and tissues metabolites, especially in terms of its broad coverage (primarily of lipids) and its amenability to quantification. It also requires very little sample volume (10 µL), is relatively inexpensive (on a per sample basis), largely automated, and offers a high-throughput route for measuring metabolites. The other methods, such as NMR and ICP–MS offer complementary data to LC–MS/MS. However, they do not provide the breadth of coverage, the level of automation or the sensitivity available via LC–MS/MS.

2.5. CONCLUSIONS

Our primary objective for undertaking these studies was to help advance the field of bovine metabolomics. This study used metabolomics techniques, including NMR, LC–MS/MS, and ICP–MS as well as literature reviews facilitated by computer-aided literature mining, to determine the number of experimentally detected or quantified metabolites in different bovine biofluids and tissues. This list of experimentally determined compounds was complemented with computer-aided genome-scale metabolite inference to provide data on thousands of other biochemically "expected" metabolites.

As reported here, the number of metabolites we experimentally detected and measured in serum, ruminal fluid, LT muscle, SM muscle, liver, and testis were 145, 145, 149, 148, 155, and

159, respectively. From this set, a total of 21 compounds or compound species (corresponding to 77 unique structures) are being reported in cattle for the first time, including 3 LysoPCs, 1 SM, 4 PCs, 10 acylcarnitines, and 3 other compounds (NADP, uridine diphosphate glucose, and uridine diphosphate-N-acetylglucosamine), all of which have been added to the BMDB. Our experimentally acquired data corresponds to a total of 3.5% (1834 out of 51,801) of the total number of bovine metabolites/chemicals reported in the BMDB. Our literature-mining efforts identified 1339 (experimentally measured) bovine metabolites, which corresponds to 2.6% of the bovine metabolome. Finally, our genome-scale inference techniques generated 48,628 biochemically expected metabolites, which accounts for the remaining 94% of the bovine metabolome in BMDB.

As far as we are aware, this compilation represents the most complete chemical inventory or chemical assessment of bovine biofluids and tissues that has been achieved to date. All of this information along with other details regarding concentration ranges, chemical structures, names, chemical classes, NMR and MS spectra, other physico-chemical properties and associated references are publicly accessible in the BMDB at www.bovinedb.ca.

Similar to previous metabolomics studies presented from our lab, we wanted to demonstrate the power and potential of quantitative metabolomics to comprehensively characterize common biofluids and tissues in cattle. The results presented here also have implications far beyond the field of metabolomics, especially given the economic importance of the bovine metabolome in the food industry and its importance in human nutrition. We expect these data to serve as a benchmark in comparing various technologies and assessing future methodological improvements in bovine metabolome research. In the meantime, it is hoped that the BMDB will provide a reliable source for metabolomics researchers, animal scientists, food

chemists, nutritional scientists, and consumers by providing a comprehensive, easy-to-use and highly centralized web-based resource on the chemical composition of bovine biofluids and tissues. The addition of more samples and the inclusion of more studies will certainly improve the quality and reliability of the data in the BMDB. Indeed, this study is most certainly not the final word on the chemical composition of the bovine metabolome. Over the coming years, we plan to further characterize the bovine biofluids and tissues metabolome using other metabolomics techniques, i.e., GC–MS, GC×GC–MS (to help identify volatile compounds), untargeted high-resolution–mass spectrometry (HR–MS), and more extensive targeted LC–MS/MS techniques to compare and extend the metabolite coverage.

2.6. SUPPLEMENTARY MATERIALS

Table 2.S1: List of rumen metabolites along with their measured or reported concentrations and their standard deviations and/or ranges (as measured in μM), Table 2.S2: List of LT and SM muscle metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g), Table 2.S3: List of liver metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g), Table 2.S4: List of testis metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g).

Table 2. 1 List of serum metabolites along with their measured or reported concentrations and their standard deviations and/or ranges (as measured in μM)

Metabolite	Platform	Concentration	Literature Value
WATER-SOLUBLE COMPOUNDS			
AMINO ACIDS			
Alanine *	LC–MS/MS and NMR	240 ± 30	151–222 [a]
Arginine	LC–MS/MS and NMR	218 ± 33	135–182 [a]
Asparagine *	LC–MS/MS and NMR	25 ± 4	20–33 [b]
Aspartate *	LC–MS/MS and NMR	24 ± 11	14–16 [c], 31–36 [a]
Beta-alanine *	NMR	8 ± 1	8–9 [d]
Citrulline *	LC–MS/MS and NMR	88 ± 15	71–84 [d]
Creatine	LC–MS/MS and NMR	196 ± 28	
Glutamate *	LC–MS/MS and NMR	92 ± 19	35–39 [d], 174–198 [a]
Glutamine *	LC–MS/MS and NMR	330 ± 42	246–260 [d]
Glycine *	LC–MS/MS and NMR	398 ± 65	405–428 [d]
Histidine *	LC–MS/MS	78 ± 10	74–84 [c]
Isoleucine	LC–MS/MS and NMR	153 ± 14	101–122 [a]
Leucine *	LC–MS/MS and NMR	212 ± 24	205–264 [c]
Lysine *	LC–MS/MS and NMR	88 ± 15	58–92 [b]
Methionine *	LC–MS/MS and NMR	34 ± 4	22–29 [c], 46–52 [a]
Ornithine *	LC–MS/MS and NMR	61 ± 13	62–135 [a]
Phenylalanine *	LC–MS/MS and NMR	71 ± 7	65–75 [c]
Proline *	LC–MS/MS and NMR	103 ± 15	84–110 [a]
Serine *	LC–MS/MS and NMR	85 ± 14	86–89 [d]
Threonine *	LC–MS/MS and NMR	74 ± 12	58–76 [c]
Tryptophan *	LC–MS/MS	47 ± 7	37–42 [c]
Tyrosine	LC–MS/MS and NMR	91 ± 10	68–75 [c]
Valine *	LC–MS/MS and NMR	356 ± 34	262–322 [c]
BIOGENIC AMINES			
Acetyl-ornithine	LC–MS/MS	3 ± 1	
Asymmetric-dimethylarginine *	LC–MS/MS	1.1 ± 0.2	1.3–2.1 [e]
Carnosine *	LC–MS/MS	30 ± 12	15–20 [d]
Creatinine *	LC–MS/MS and NMR	113 ± 17	109–140 [f]
Kynurenine *	LC–MS/MS	7 ± 2	4–7 [b]
Methionine-sulfoxide	LC–MS/MS	1.2 ± 0.3	
Methylhistidine *	LC–MS/MS	14 ± 2	2–12 [g]
Putrescine	LC–MS/MS	0.04 ± 0.02	
Sarcosine	LC–MS/MS and NMR	3 ± 1	10–12 [d]
Serotonin *	LC–MS/MS	9 ± 3	4–13 [h]
Spermidine	LC–MS/MS	0.2 ± 0.1	
Spermine	LC–MS/MS	0.2 ± 0.2	
Taurine	LC–MS/MS and NMR	80 ± 20	33–47 [d]
Total-dimethylarginine	LC–MS/MS	2.1 ± 0.3	
Trans-hydroxyproline	LC–MS/MS	25 ± 5	
Trimethylamine N-oxide	LC–MS/MS	6 ± 3	
CARBOHYDRATES			
Glucose *	LC–MS/MS and NMR	3962 ± 443	3290–4070 [i]
ORGANIC ACIDS			
3-hydroxybutyrate *	NMR	340 ± 145	250–2110 [j]
Acetate	NMR	403 ± 199	920–1040 [k]
Alpha-aminoadipate	LC–MS/MS	1.28 ± 0.52	7.3–8.1 [d]
Ascorbate (Vitamin C) *	NMR	11 ± 3	8–18 [l]
Formate	NMR	78 ± 12	
Fumarate	NMR	1.2 ± 0.2	
Lactate	NMR	4850 ± 2017	658–1600 [m]

Pyruvate	NMR	150 ± 40	
MISCELANEOUS			
Acetone [*]	NMR	70 ± 22	80–990 [j]
Betaine	LC–MS/MS and NMR	169 ± 31	14–26 [n]
Choline	LC–MS/MS and NMR	20 ± 4	4–5 [n]
Ethanol [*]	NMR	8 ± 1	3–68 [i]
Glycerol	NMR	314 ± 38	
Isopropanol	NMR	2 ± 1	
Methanol	NMR	32 ± 4	
Myo-inositol	NMR	45 ± 11	
Urea	NMR	1321 ± 282	1950–4080 [o]
Uridine	NMR	3 ± 1	
LIPID-LIKE COMPOUNDS			
PHOSPHATIDYLCHOLINES, ACYL-ALKYL			
PC ae (36:0)	LC–MS/MS	1.6 ± 0.4	
PC ae (40:6)	LC–MS/MS	0.46 ± 0.11	
PHOSPHATIDYLCHOLINES, DIACYL			
PC aa (32:2)	LC–MS/MS	4 ± 1	
PC aa (36:6)	LC–MS/MS	0.7 ± 0.2	
PC aa (36:0)	LC–MS/MS	6 ± 2	
PC aa (38:6)	LC–MS/MS	1 ± 0.3	
PC aa (38:0)	LC–MS/MS	0.8 ± 0.2	
PC aa (40:6)	LC–MS/MS	1.6 ± 0.4	
PC aa (40:2)	LC–MS/MS	0.4 ± 0.1	
PC aa (40:1)	LC–MS/MS	0.21 ± 0.04	
LYSOPHOSPHATIDYLCHOLINES, ACYL C			
LysoPC(14:0)	LC–MS/MS	0.8 ± 0.1	
LysoPC(16:1)	LC–MS/MS	0.6 ± 0.1	
LysoPC(16:0) [*]	LC–MS/MS	20 ± 4	15–58 [n]
LysoPC(17:0)	LC–MS/MS	3 ± 1	
LysoPC(18:2)	LC–MS/MS	15 ± 3	30–186 [n]
LysoPC(18:1)	LC–MS/MS	6 ± 1	18–69 [n]
LysoPC(18:0) [*]	LC–MS/MS	30 ± 5	14–82 [n]
LysoPC(20:4)	LC–MS/MS	0.48 ± 0.14	
LysoPC(20:3)	LC–MS/MS	1.6 ± 0.3	
LysoPC(24:0)	LC–MS/MS	0.051 ± 0.012	
LysoPC(26:1)	LC–MS/MS	0.1 ± 0.04	
LysoPC(26:0)	LC–MS/MS	0.7 ± 0.3	
LysoPC(28:1)	LC–MS/MS	0.3 ± 0.1	
LysoPC(28:0)	LC–MS/MS	0.28 ± 0.11	
SPHINGOMYELINS			
SM(16:1)	LC–MS/MS	5 ± 1	
SM(16:0)	LC–MS/MS	68 ± 10	
SM(18:1)	LC–MS/MS	11 ± 3	
SM(18:0)	LC–MS/MS	12 ± 2	
SM(20:2)	LC–MS/MS	1.1 ± 0.3	
HYDROXYSPHINGOMYELINS			
SM(14:1(OH))	LC–MS/MS	5 ± 1	
SM(16:1(OH))	LC–MS/MS	9 ± 2	
SM(22:2(OH))	LC–MS/MS	4 ± 1	
SM(22:1(OH))	LC–MS/MS	9 ± 1	
SM(24:1(OH))	LC–MS/MS	2 ± 0.4	
ACYLCARNITINES			
C0 (Carnitine)	LC–MS/MS	7 ± 1	
C2 (Acetylcarnitine) [*]	LC–MS/MS	2 ± 1	0.65–1.09 [b]
C3:1 (Propenoylcarnitine)	LC–MS/MS	0.029 ± 0.004	

C3 (Propionylcarnitine)	LC–MS/MS	0.2 ± 0.04	
C4:1 (Butenylcarnitine)	LC–MS/MS	0.017 ± 0.002	
C4 (Butyrylcarnitine)	LC–MS/MS	0.2 ± 0.1	
C3-OH (Hydroxypropionylcarnitine) *	LC–MS/MS	0.027 ± 0.004	0.01–0.02 [b]
C5:1 (Tiglylcarnitine)	LC–MS/MS	0.023 ± 0.004	
C5 (Valerylcarnitine) *	LC–MS/MS	0.09 ± 0.03	0.03–0.06 [b]
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC–MS/MS	0.04 ± 0.01	
C6:1 (Hexenoylcarnitine)	LC–MS/MS	0.02 ± 0.01	
C6 (C4:1-DC) (Hexanoylcarnitine)	LC–MS/MS	0.05 ± 0.01	0.02–0.03 [b]
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine) *	LC–MS/MS	0.04 ± 0.01	0.05–0.06 [b]
C5:1-DC (Glutaconylcarnitine)	LC–MS/MS	0.018 ± 0.003	
C5-DC (C6-OH)(Glutarylcarnitine)	LC–MS/MS	0.03 ± 0.01	
C8 (Octanoylcarnitine)	LC–MS/MS	0.02 ± 0.01	
C5-M-DC (methylglutarylcarnitine)	LC–MS/MS	0.0196 ± 0.0024	
C9 (Nonaylcarnitine)	LC–MS/MS	0.022 ± 0.003	
C7-DC (Pimelylcarnitine) *	LC–MS/MS	0.04 ± 0.04	0.01–0.02 [b]
C10:2 (Decadienylcarnitine)	LC–MS/MS	0.06 ± 0.01	
C10:1 (Decenoylcarnitine)	LC–MS/MS	0.17 ± 0.03	
C10 (Decanoylcarnitine)	LC–MS/MS	0.18 ± 0.04	
C12:1 (Dodecenoylcarnitine)	LC–MS/MS	0.084 ± 0.014	
C12 (Dodecanoylcarnitine) *	LC–MS/MS	0.04 ± 0.01	0.02–0.03 [b]
C14:2 (Tetradecadienylcarnitine)	LC–MS/MS	0.03 ± 0.01	
C14:1 (Tetradecenoylcarnitine)	LC–MS/MS	0.0518 ± 0.0103	
C14 (Tetradecanoylcarnitine) *	LC–MS/MS	0.02 ± 0.01	0.01–0.02 [b]
C12-DC (Dodecanedioylcarnitine)	LC–MS/MS	0.018 ± 0.003	
C14:2-OH (Hydroxytetradecadienylcarnitine)	LC–MS/MS	0.008 ± 0.002	
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002	
C16:2 (Hexadecadienylcarnitine)	LC–MS/MS	0.012 ± 0.002	
C16:1 (Hexadecenoylcarnitine)	LC–MS/MS	0.029 ± 0.003	
C16 (Hexadecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01	
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC–MS/MS	0.005 ± 0.001	
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC–MS/MS	0.0184 ± 0.0034	
C16-OH (Hydroxyhexadecanoylcarnitine) *	LC–MS/MS	0.008 ± 0.001	0.003–0.006 [b]
C18:2 (Octadecadienylcarnitine)	LC–MS/MS	0.007 ± 0.001	
C18:1 (Octadecenoylcarnitine)	LC–MS/MS	0.0147 ± 0.0031	
C18 (Octadecanoylcarnitine)	LC–MS/MS	0.021 ± 0.008	
C18:1-OH (Hydroxyoctadecenoylcarnitine) *	LC–MS/MS	0.009 ± 0.001	0.008–0.009 [b]
TRACE ELEMENTAL COMPOUNDS			
Sodium *	ICP–MS	133515 ± 13658	107400–108600 [p], 136000–136710 [q]
Magnesium *	ICP–MS	931 ± 88	850–920 [f]
Phosphorus *	ICP–MS	1298 ± 164	1350–1620 [p]
Potassium *	ICP–MS	4296 ± 388	4060–4340 [f]
Calcium *	ICP–MS	2228 ± 221	1400–2200 [h]
Iron *	ICP–MS	52 ± 13	50–51 [r]
Copper *	ICP–MS	9 ± 2	6–9 [r]
Zinc *	ICP–MS	12 ± 2	14–18 [r]
Selenium *	ICP–MS	1.4 ± 0.2	0.5–2.7 [s]
Rubidium	ICP–MS	1.8 ± 0.2	
Strontium	ICP–MS	1 ± 0.1	
Cesium	ICP–MS	0.0017 ± 0.0003	
Barium	ICP–MS	0.2 ± 0.03	

[*] Compounds that exhibited good agreement with literature values; [a] Motyl and Barej, 1986 [40]; [b] Sadri et al., 2017 [41]; [c] Greenwood et al., 2001 [42]; [d] Zhou et al., 2016 [37]; [e] Chan et al., 2000 [43]; [f] Consolo et al., 2018 [44]; [g] van der Drift et al., 2012 [45]; [h] Hernandez-Castellano et al., 2017 [46]; [I] Raun and Kristensen, 2011 [33]; [j] Sato, 2009 [34]; [k] Sato et al., 1999 [35]; [l] Padilla et al., 2006 [47]; [m] Kenny et al., 2002 [48]; [n] Artegoitia et al., 2014 [49]; [o] Liker et al., 2005 [50]; [p] Nozad et al., 2012 [51]; [q] Macdonald et al., 2017 [52]; [r] Noaman et al., 2012 [30]; [s] Waldner et al., 1998 [53].

Table 2. 2 Metabolome coverage of different bovine biofluids and tissues in the BMDB

Tissue/Biofluid Location	Identified Metabolites or Metabolite Species	Identified Metabolites with Unique Structures	Quantified Metabolites with Unique Structures
BIOFLUID			
Blood	330	453	296
Colostrum	70	70	4
Milk	928	2350	1652
Ruminal fluid	328	769	728
Semen	76	76	0
Urine	62	62	0
TISSUE			
Adipose tissue	199	199	71
Brain	557	1887	0
Epidermis	275	275	0
Fibroblasts	327	327	0
Intestine	253	253	0
Kidney	531	615	0
Liver	1056	1254	273
Longissimus thoracis muscle	153	267	267
Mammary gland	269	269	0
Neuron	322	322	0
Pancreas	114	114	0
Placenta	579	586	0
Platelet	204	204	0
Prostate	268	268	0
Semimembranosus muscle	153	267	267
Skeletal muscle	382	496	274
Spleen	168	168	0
Testis	328	442	277
All tissues	**857**	**4464**	**N/A** *

Note: Metabolite species refer to those molecules with non-unique chemical formulas or masses (such as lipid isomers), while metabolites with unique structures correspond to compounds with a unique and clearly defined chemical structure and a unique chemical name; * Not available.

Table 2. 3 Comparison of concentrations of 26 metabolites between nuclear magnetic resonance (NMR) and liquid chromatography–tandem mass spectrometry (LC–MS/MS) methods along with their measured or reported concentrations and their standard deviations and/or ranges in bovine serum (as measured in μM)

Compound Name	NMR	LC–MS/MS	Average Difference (%)
Alanine	252 ± 31	240 ± 30	4
Arginine	210 ± 28	218 ± 33	3
Asparagine	26 ± 5	25 ± 4	3
Aspartate	Glucose overlap	24 ± 11	
Betaine	180 ± 37	169 ± 31	6
Choline	19 ± 4	20 ± 4	5
Citrulline	89 ± 16	88 ± 15	1
Creatine	210 ± 30	196 ± 28	6
Creatinine	121 ± 17	113 ± 17	6
Glucose	4572 ± 588	3962 ± 443	14
Glutamate	Proline overlap	92 ± 19	
Glutamine	360 ± 47	330 ± 42	8
Glycine	438 ± 71	398 ± 65	9
Isoleucine	160 ± 22	153 ± 14	4
Leucine	216 ± 30	212 ± 24	1
Lysine	Arginine overlap	88 ± 15	
Methionine	35 ± 5	34 ± 4	2
Ornithine	67 ± 11	61 ± 13	9
Phenylalanine	67 ± 9	71 ± 7	5
Proline	94 ± 16	103 ± 15	9
Sarcosine	4 ± 1	3 ± 1	28
Serine	87 ± 18	85 ± 14	2
Taurine	Glucose overlap	80 ± 20	
Threonine	72 ± 11	74 ± 12	2
Tyrosine	84 ± 11	91 ± 10	8
Valine	390 ± 48	356 ± 34	9

Table 2. S1 List of rumen metabolites along with their measured or reported concentrations and their standard deviations and/or ranges (as measured in μM)

Metabolite	Platform	Concentration	Literature value
WATER-SOLUBLE COMPOUNDS			
AMINO ACIDS			
Alanine *	LC–MS/MS & NMR	1060 ± 484	83–799 [a]
Arginine *	LC–MS/MS	17 ± 11	5–81 [b]
Asparagine	LC–MS/MS	1 ± 1	
Aspartate *	LC–MS/MS & NMR	566 ± 204	153–925 [a]
Beta-alanine *	NMR	72 ± 38	3–33 [c]
Citrulline	LC–MS/MS	184 ± 74	
Creatine *	LC–MS/MS & NMR	3 ± 4	2–15 [c]
Glutamate *	LC–MS/MS & NMR	1702 ± 471	238–887 [a]
Glycine *	LC–MS/MS & NMR	974 ± 315	50–696 [a]
Histidine *	LC–MS/MS	44 ± 28	18–69 [a]
Isoleucine *	LC–MS/MS & NMR	556 ± 127	123–1210 [a]
Leucine *	LC–MS/MS & NMR	603 ± 119	76–571 [a]
Lysine *	LC–MS/MS & NMR	2085 ± 869	91–1095 [a]
Methionine	LC–MS/MS & NMR	268 ± 102	9–66 [b]
Ornithine	LC–MS/MS	149 ± 77	7–51 [b]
Phenylalanine	LC–MS/MS & NMR	310 ± 129	17–85 [b]
Proline *	LC–MS/MS & NMR	834 ± 377	240–1275 [a]
Serine	LC–MS/MS & NMR	547 ± 187	24–180 [b]
Threonine	LC–MS/MS & NMR	658 ± 301	53–153 [c]
Tryptophan *	LC–MS/MS & NMR	24 ± 14	4–26 [b]
Tyrosine *	LC–MS/MS & NMR	295 ± 136	68–471 [a]
Valine *	LC–MS/MS & NMR	848 ± 295	71–593 [a]
BIOGENIC AMINES			
Acetyl-ornithine	LC–MS/MS	9 ± 5	0.1–2.5 [b]
Asymmetric-dimethylarginine *	LC–MS/MS	4 ± 3	0.19–1 [b]
Total-dimethylarginine	LC–MS/MS	5 ± 3	19–47 [b]
Histamine *	LC–MS/MS	57 ± 43	1–12 [b]
Methionine-sulfoxide *	LC–MS/MS	26 ± 9	5–37 [b]
Methylamine *	NMR	132 ± 116	27–822 [b]
Methylhistidine	LC–MS/MS	2 ± 1	
Putrescine *	LC–MS/MS & NMR	103 ± 58	16–303 [b]
Sarcosine	LC–MS/MS & NMR	2 ± 1	6–67 [b]
Serotonin	LC–MS/MS	0.1 ± 0.1	0.9–1.1 [b]
Spermidine	LC–MS/MS	37 ± 22	
Spermine	LC–MS/MS	3 ± 3	
Taurine *	LC–MS/MS & NMR	8 ± 7	1–2 [b]
Trans-hydroxyproline	LC–MS/MS	2 ± 1	
CARBOHYDRATES			
Glucose	LC–MS/MS & NMR	15934 ± 11188	393–3111 [b]

Ribose	NMR	886 ± 201	177–501 [b]
ORGANIC ACIDS			
2-oxoglutarate	NMR	40 ± 24	
3-phenylpropionate	NMR	51 ± 18	296–713 [b]
Acetate [*]	NMR	37264 ± 7582	41000–81000 [a]
Alpha-aminoadipate	LC–MS/MS	1 ± 1	
Butyrate [*]	NMR	25848 ± 9273	6472–18088 [b]
Formate	NMR	20 ± 4	60–655 [a]
Fumarate [*]	NMR	18 ± 7	19–315 [a]
Isobutyrate [*]	NMR	719 ± 226	708–1100 [c]
Lactate [*]	NMR	2029 ± 1165	224–1560 [a]
Nicotinate [*]	NMR	60 ± 10	29–79 [b]
Phenylacetate [*]	NMR	189 ± 57	212–785 [a]
Propionate [*]	NMR	16103 ± 4550	14000–17000 [a]
Pyroglutamate [*]	NMR	233 ± 64	311–645 [b]
Pyruvate	NMR	62 ± 33	
Succinate [*]	NMR	597 ± 344	40–289 [a]
Valerate [*]	NMR	2739 ± 677	1500–5000 [a]
MISCELANEOUS			
Acetone	NMR	66 ± 40	5–19 [c]
Betaine [*]	LC–MS/MS	15 ± 40	2–17 [c]
Choline	LC–MS/MS & NMR	105 ± 59	4–40 [c]
Ethanol [*]	NMR	556 ± 450	600–3200 [d]
Glycerol	NMR	685 ± 231	89–336 [b]
Hypoxanthine	NMR	753 ± 191	38–266 [b]
Isopropanol [*]	NMR	103 ± 45	14–70 [c]
Thymine	NMR	188 ± 45	15–63 [b]
Uracil	NMR	1007 ± 230	109–405 [c]
Uridine [*]	NMR	18 ± 10	2–17 [c]
LIPID-LIKE COMPOUNDS			
PHOSPHATIDYLCHOLINES, ACYL-ALKYL			
PC ae (36:0)	LC–MS/MS	0.05 ± 0.04	
PC ae (40:6)	LC–MS/MS	0.04 ± 0.01	
PHOSPHATIDYLCHOLINES, DIACYL			
PC aa (32:2)	LC–MS/MS	0.048 ± 0.024	
PC aa (36:0)	LC–MS/MS	0.07 ± 0.04	
PC aa (38:6)	LC–MS/MS	0.06 ± 0.04	
PC aa (38:0)	LC–MS/MS	0.08 ± 0.01	
PC aa (40:6)	LC–MS/MS	0.03 ± 0.01	
PC aa (40:1)	LC–MS/MS	0.016 ± 0.004	
LYSOPHOSPHATIDYLCHOLINES, ACYL C			
LysoPC(14:0)	LC–MS/MS	6 ± 3	
LysoPC(16:0) [*]	LC–MS/MS	0.3 ± 0.2	0.01–0.2 [b]
LysoPC(18:2)	LC–MS/MS	0.2 ± 0.1	
LysoPC(18:1)	LC–MS/MS	0.074 ± 0.051	

LysoPC(18:0)	LC–MS/MS	0.2 ± 0.1	
LysoPC(20:4)	LC–MS/MS	0.03 ± 0.02	
LysoPC(20:3)	LC–MS/MS	2 ± 1	
LysoPC(24:0)	LC–MS/MS	0.161 ± 0.153	
LysoPC(26:0)	LC–MS/MS	0.7 ± 0.3	
LysoPC(28:0)	LC–MS/MS	0.11 ± 0.03	
SPHINGOMYELINS			
SM(16:0)	LC–MS/MS	0.75 ± 0.62	0.01–0.05 [b]
SM(18:1)	LC–MS/MS	0.04 ± 0.03	
SM(18:0)	LC–MS/MS	0.19 ± 0.13	
HYDROXYSPHINGOMYELINS			
SM(14:1(OH))	LC–MS/MS	0.05 ± 0.01	0.01–0.03 [b]
SM(16:1(OH))	LC–MS/MS	0.07 ± 0.04	0.01–0.02 [b]
SM(22:1(OH)) [*]	LC–MS/MS	0.1 ± 0.1	0.01–0.1 [b]
ACYLCARNITINES			
C0 (Carnitine)	LC–MS/MS	3 ± 1	
C2 (Acetylcarnitine)	LC–MS/MS	0.24 ± 0.11	
C3:1 (Propenoylcarnitine)	LC–MS/MS	0.03 ± 0.01	
C3 (Propionylcarnitine)	LC–MS/MS	0.09 ± 0.02	0.03–0.05 [b]
C4:1 (Butenylcarnitine) [*]	LC–MS/MS	0.04 ± 0.01	0.03–0.04 [b]
C4 (Butyrylcarnitine)	LC–MS/MS	0.05 ± 0.02	
C3-OH (Hydroxypropionylcarnitine) [*]	LC–MS/MS	0.04 ± 0.01	0.03–0.04 [b]
C5:1 (Tiglylcarnitine) [*]	LC–MS/MS	0.03 ± 0.01	0.04–0.07 [b]
C5 (Valerylcarnitine) [*]	LC–MS/MS	0.03 ± 0.01	0.03–0.06 [b]
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC–MS/MS	1 ± 1	0.04–0.33 [b]
C6:1 (Hexenoylcarnitine)	LC–MS/MS	0.04 ± 0.01	0.09–0.10 [b]
C6 (C4:1-DC) (Hexanoylcarnitine)	LC–MS/MS	0.08 ± 0.01	
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine)	LC–MS/MS	0.027 ± 0.004	
C5:1-DC (Glutaconylcarnitine)	LC–MS/MS	0.019 ± 0.003	0.039–0.045 [b]
C5-DC (C6-OH)(Glutarylcarnitine)	LC–MS/MS	0.015 ± 0.003	
C8 (Octanoylcarnitine)	LC–MS/MS	0.014 ± 0.003	
C5-M-DC (methylglutarylcarnitine)	LC–MS/MS	0.022 ± 0.003	0.174–0.2 [b]
C9 (Nonaylcarnitine)	LC–MS/MS	0.011 ± 0.002	
C7-DC (Pimelylcarnitine) [*]	LC–MS/MS	0.029 ± 0.021	0.03–0.07 [b]
C10:2 (Decadienylcarnitine)	LC–MS/MS	0.06 ± 0.01	
C10:1 (Decenoylcarnitine)	LC–MS/MS	0.2 ± 0.1	
C10 (Decanoylcarnitine)	LC–MS/MS	0.1 ± 0.02	
C12:1 (Dodecenoylcarnitine)	LC–MS/MS	0.1 ± 0.02	
C12 (Dodecanoylcarnitine)	LC–MS/MS	0.03 ± 0.01	
C14:2 (Tetradecadienylcarnitine) [*]	LC–MS/MS	0.023 ± 0.004	0.017–0.023 [b]
C14:1 (Tetradecenoylcarnitine)	LC–MS/MS	0.007 ± 0.002	0.089–0.091 [b]
C14 (Tetradecanoylcarnitine)	LC–MS/MS	0.009 ± 0.002	
C12-DC (Dodecanedioylcarnitine)	LC–MS/MS	0.037 ± 0.012	
C14:2-OH (Hydroxytetradecadienylcarnitine) [*]	LC–MS/MS	0.011 ± 0.003	0.008–0.032 [b]
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002	

C16:2 (Hexadecadienylcarnitine) *	LC–MS/MS	0.008 ± 0.001	0.002–0.013 [b]
C16:1 (Hexadecenoylcarnitine)	LC–MS/MS	0.023 ± 0.004	
C16 (Hexadecanoylcarnitine) *	LC–MS/MS	0.016 ± 0.004	0.01–0.055 [b]
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC–MS/MS	0.009 ± 0.002	
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC–MS/MS	0.01 ± 0.002	
C16-OH (Hydroxyhexadecanoylcarnitine) *	LC–MS/MS	0.012 ± 0.003	0.009–0.011 [b]
C18:2 (Octadecadienylcarnitine)	LC–MS/MS	0.006 ± 0.001	0.07–0.072 [b]
C18:1 (Octadecenoylcarnitine)	LC–MS/MS	0.015 ± 0.003	
C18 (Octadecanoylcarnitine)	LC–MS/MS	0.0067 ± 0.0021	
C18:1-OH (Hydroxyoctadecenoylcarnitine) *	LC–MS/MS	0.012 ± 0.003	0.002–0.022 [b]
TRACE ELEMENTAL COMPOUNDS			
Lithium	ICP–MS	21 ± 5	3–4 [b]
Sodium	ICP–MS	235634 ± 19788	110000–117000 [b]
Magnesium	ICP–MS	7465 ± 3511	96–108 [b]
Phosphorus	ICP–MS	12395 ± 1908	9140–9270 [b]
Potassium	ICP–MS	39878 ± 6579	17980–18270 [b]
Calcium *	ICP–MS	371 ± 722	904–958 [b]
Titanium	ICP–MS	2.5 ± 0.4	
Manganese *	ICP–MS	59 ± 62	2–3 [b]
Iron *	ICP–MS	40 ± 8	21–32 [b]
Nickel	ICP–MS	2 ± 1	0.04–0.2 [b]
Cobalt	ICP–MS	1 ± 0.2	
Copper	ICP–MS	5 ± 1	2–3 [b]
Zinc	ICP–MS	10 ± 3	2–3 [b]
Rubidium	ICP–MS	25 ± 5	5–6 [b]
Strontium *	ICP–MS	4 ± 2	1–2 [b]
Cesium	ICP–MS	0.03 ± 0.01	0.001–0.01 [b]
Barium *	ICP–MS	2 ± 1	0.5–1 [b]

* Compounds that exhibited good agreement with literature values; [a] Lee et al., 2012 [197]; [b] Saleem et al., 2013 [198]; [c] O'Callaghan et al., 2018 [199]; [d] Raun and Kristensen, 2011 [200].

Table 2. S2 List of LT and SM muscle metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g)

Metabolite	Platform	Concentration (LT muscle)	Concentration (SM muscle)	Literature value
WATER-SOLUBLE COMPOUNDS				
AMINO ACIDS				
Alanine *	LC–MS/MS & NMR	1465 ± 272	1472 ± 278	1055–1310 [a]
Arginine	LC–MS/MS	61 ± 14	70 ± 16	299–686 [b]
Asparagine	LC–MS/MS	40 ± 12	46 ± 17	106–225 [b]
Aspartate *	LC–MS/MS & NMR	70 ± 28	57 ± 28	14–54 [b]
Beta-alanine *	NMR	149 ± 57	122 ± 37	84–155 [b]
Citrulline	LC–MS/MS	34 ± 12	31 ± 14	59–111 [b]
Creatine *	LC–MS/MS & NMR	4672 ± 438	4755 ± 306	1040–1520 [c], 7262–10319 [a]
Glutamate *	LC–MS/MS & NMR	425 ± 180	541 ± 234	134–942 [b]
Glutamine *	LC–MS/MS & NMR	2841 ± 828	2492 ± 668	3051–4834 [b]
Glutathione	NMR	226 ± 51	162 ± 46	769–775 [b]
Glycine *	LC–MS/MS & NMR	576 ± 215	538 ± 147	342–570 [a]
Histidine	LC–MS/MS	75 ± 18	85 ± 20	134–281 [b]
Isoleucine *	LC–MS/MS & NMR	107 ± 22	116 ± 23	43–231 [a]
Leucine *	LC–MS/MS & NMR	154 ± 37	158 ± 33	80–424 [a]
Lysine	LC–MS/MS	59 ± 18	66 ± 19	306–745 [b]
Methionine *	LC–MS/MS & NMR	34 ± 10	37 ± 11	40–395 [b]
Ornithine	LC–MS/MS	35 ± 12	40 ± 10	80–187 [b]
Phenylalanine *	LC–MS/MS & NMR	55 ± 12	58 ± 10	51–258 [a]
Proline	LC–MS/MS	174 ± 59	203 ± 60	240–365 [b]
Serine	LC–MS/MS	165 ± 36	161 ± 31	273–830 [b]
Threonine *	LC–MS/MS & NMR	194 ± 80	193 ± 32	193–468 [b]
Tryptophan	LC–MS/MS & NMR	19 ± 5	21 ± 3	35–95 [b]
Tyrosine *	LC–MS/MS & NMR	54 ± 14	52 ± 10	44–269 [a]
Valine *	LC–MS/MS & NMR	270 ± 62	278 ± 49	71–288 [a]
BIOGENIC AMINES				
Acetyl-ornithine	LC–MS/MS	18 ± 8	21 ± 8	
Asymmetric-dimethylarginine	LC–MS/MS	0.8 ± 0.2	0.9 ± 0.2	
Carnosine	LC–MS/MS & NMR	22085 ± 4859	21958 ± 3048	9690–13658 [b]
Creatinine *	LC–MS/MS & NMR	313 ± 48	315 ± 47	128–429 [a]
Histamine	LC–MS/MS	17 ± 4	23 ± 6	
Kynurenine	LC–MS/MS	0.6 ± 0.2	0.5 ± 0.2	
Methionine-sulfoxide	LC–MS/MS	3 ± 7	2 ± 8	
Methylhistidine	LC–MS/MS	35 ± 8	35 ± 9	
Putrescine	LC–MS/MS	1.3 ± 0.4	2 ± 1	9–22 [b]
Sarcosine	LC–MS/MS & NMR	12 ± 5	10 ± 4	
Spermidine	LC–MS/MS	0.09 ± 0.03	0.11 ± 0.04	
Spermine	LC–MS/MS	0.2 ± 0.1	0.11 ± 0.03	

Taurine	LC–MS/MS & NMR	480 ± 146	844 ± 301	
Total-dimethylarginine	LC–MS/MS	1.7 ± 0.4	2 ± 1	
Trans-hydroxyproline	LC–MS/MS	32 ± 11	31 ± 11	
Trimethylamine N-oxide	LC–MS/MS	14 ± 5	14 ± 4	
CARBOHYDRATES				
Glucose-1-phosphate *	NMR	184 ± 58	175 ± 48	120 [d], 363–1178 [b]
Glucose	LC–MS/MS & NMR	536 ± 259	754 ± 403	3300 [d]
ORGANIC ACIDS				
3-hydroxybutyrate *	NMR	97 ± 27	88 ± 33	114–242 [b]
Acetate *	NMR	188 ± 31	219 ± 52	97–291 [a]
Alpha-aminoadipate	LC–MS/MS	7 ± 2	10 ± 4	
Formate	NMR	533 ± 316	723 ± 184	110 [d]
Fumarate *	NMR	141 ± 64	220 ± 76	36–208 [b]
Inosinate (IMP) *	NMR	15 ± 11	124 ± 136	31–59 [c], 40–8977 [b]
Lactate *	NMR	31131 ± 8268	37879 ± 10580	14262–21850 [a], 37175–131553 [b]
Nicotinurate	NMR	96 ± 38	170 ± 54	
Pyruvate *	NMR	123 ± 51	157 ± 68	1–183 [b]
Succinate *	NMR	1133 ± 308	743 ± 384	891–2314 [b]
MISCELANEOUS				
Betaine *	LC–MS/MS & NMR	1321 ± 351	1139 ± 246	1477–1826 [b]
Choline	LC–MS/MS & NMR	15 ± 7	15 ± 6	34–153 [b]
Ethanol *	NMR	299 ± 182	246 ± 172	160 [d]
Glycerol *	NMR	328 ± 210	326 ± 174	392–528 [c]
Myo-inositol	NMR	504 ± 133	593 ± 126	
NAD+ *	NMR	641 ± 127	515 ± 70	15–611 [b]
NADH	NMR	94 ± 75	101 ± 45	
O-acetylcarnitine	NMR	1892 ± 398	1612 ± 318	
LIPID-LIKE COMPOUNDS				
PHOSPHATIDYLCHOLINES, ACYL-ALKYL				
PC ae (36:0)	LC–MS/MS	0.55 ± 0.12	0.6 ± 0.2	
PC ae (40:6)	LC–MS/MS	0.44 ± 0.12	0.56 ± 0.14	
PHOSPHATIDYLCHOLINES, DIACYL				
PC aa (32:2)	LC–MS/MS	1.7 ± 0.4	3 ± 2	
PC aa (36:6)	LC–MS/MS	0.3 ± 0.1	0.3 ± 0.1	
PC aa (36:0)	LC–MS/MS	0.85 ± 0.14	1 ± 0.2	
PC aa (38:6)	LC–MS/MS	0.5 ± 0.1	0.6 ± 0.2	
PC aa (38:0)	LC–MS/MS	0.4 ± 0.1	0.4 ± 0.1	
PC aa (40:6)	LC–MS/MS	0.13 ± 0.03	0.14 ± 0.04	
PC aa (40:2)	LC–MS/MS	0.017 ± 0.004	0.02 ± 0.01	
PC aa (40:1)	LC–MS/MS	0.026 ± 4	0.031 ± 0.004	
LYSOPHOSPHATIDYLCHOLINES, ACYL C				
LysoPC(14:0)	LC–MS/MS	0.07 ± 0.01	0.07 ± 0.01	

LysoPC(16:1)	LC–MS/MS	0.02 ± 0.01	0.03 ± 0.01	
LysoPC(16:0)	LC–MS/MS	0.1 ± 0.02	0.11 ± 0.03	
LysoPC(17:0)	LC–MS/MS	0.013 ± 0.003	0.016 ± 0.004	
LysoPC(18:2)	LC–MS/MS	0.3 ± 0.1	0.4 ± 0.1	
LysoPC(18:1)	LC–MS/MS	0.3 ± 0.1	0.4 ± 0.1	
LysoPC(18:0)	LC–MS/MS	0.07 ± 0.02	0.08 ± 0.03	
LysoPC(20:4)	LC–MS/MS	0.07 ± 0.03	0.09 ± 0.03	
LysoPC(20:3)	LC–MS/MS	0.05 ± 0.01	0.08 ± 0.02	
LysoPC(24:0)	LC–MS/MS	0.023 ± 0.004	0.03 ± 0.01	
LysoPC(26:1)	LC–MS/MS	0.05 ± 0.01	0.06 ± 0.02	
LysoPC(26:0)	LC–MS/MS	0.05 ± 0.01	0.06 ± 0.02	
LysoPC(28:1)	LC–MS/MS	0.04 ± 0.01	0.04 ± 0.01	
LysoPC(28:0)	LC–MS/MS	0.05 ± 0.01	0.06 ± 0.02	
SPHINGOMYELINS				
SM(16:1)	LC–MS/MS	0.11 ± 0.02	0.13 ± 0.03	
SM(16:0)	LC–MS/MS	3 ± 1	2.45 ± 0.71	
SM(18:1)	LC–MS/MS	1.1 ± 0.2	1.2 ± 0.4	
SM(18:0)	LC–MS/MS	8 ± 1	8 ± 2	
SM(20:2)	LC–MS/MS	0.4 ± 0.1	0.5 ± 0.2	
HYDROXYSPHINGOMYELINS				
SM(14:1(OH))	LC–MS/MS	0.2 ± 0.04	0.2 ± 0.1	
SM(16:1(OH))	LC–MS/MS	1.1 ± 0.2	1.2 ± 0.4	
SM(22:2(OH))	LC–MS/MS	0.4 ± 0.1	0.47 ± 0.1	
SM(22:1(OH))	LC–MS/MS	0.9 ± 0.2	0.94 ± 0.22	
SM(24:1(OH))	LC–MS/MS	0.08 ± 0.02	0.09 ± 0.03	
ACYLCARNITINES				
C0 (Carnitine) [*]	LC–MS/MS	1856 ± 424	1751 ± 337	791–1143 [a], 2600 [d]
C2 (Acetylcarnitine)	LC–MS/MS	854 ± 153	871 ± 150	
C3:1 (Propenoylcarnitine)	LC–MS/MS	1.8 ± 0.2	1.8 ± 0.2	
C3 (Propionylcarnitine)	LC–MS/MS	6 ± 2	7 ± 2	
C4:1 (Butenylcarnitine)	LC–MS/MS	0.07 ± 0.01	0.08 ± 0.03	
C4 (Butyrylcarnitine)	LC–MS/MS	16 ± 6	22 ± 8	
C3-OH (Hydroxypropionylcarnitine)	LC–MS/MS	0.2 ± 0.1	0.2 ± 0.1	
C5:1 (Tiglylcarnitine)	LC–MS/MS	0.3 ± 0.1	0.3 ± 0.1	
C5 (Valerylcarnitine)	LC–MS/MS	4 ± 2	4 ± 2	
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC–MS/MS	6 ± 3	7 ± 5	
C6:1 (Hexenoylcarnitine)	LC–MS/MS	0.12 ± 0.03	0.2 ± 0.1	
C6 (C4:1-DC) (Hexanoylcarnitine)	LC–MS/MS	0.52 ± 0.51	0.7 ± 0.4	
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine)	LC–MS/MS	1.6 ± 0.4	1.7 ± 0.4	
C5:1-DC (Glutaconylcarnitine)	LC–MS/MS	0.04 ± 0.03	0.05 ± 0.03	
C5-DC (C6-OH)(Glutarylcarnitine)	LC–MS/MS	0.09 ± 0.04	0.14 ± 0.11	
C8 (Octanoylcarnitine)	LC–MS/MS	0.26 ± 0.32	0.27 ± 0.24	
C5-M-DC (methylglutarylcarnitine)	LC–MS/MS	0.02 ± 0.01	0.02 ± 0.01	
C9 (Nonaylcarnitine)	LC–MS/MS	0.03 ± 0.02	0.04 ± 0.03	

C7-DC (Pimelylcarnitine)	LC–MS/MS	0.1 ± 0.1	0.14 ± 0.11
C10:2 (Decadienylcarnitine)	LC–MS/MS	0.04 ± 0.01	0.024 ± 0.004
C10:1 (Decenoylcarnitine)	LC–MS/MS	0.3 ± 0.1	0.3 ± 0.1
C10 (Decanoylcarnitine)	LC–MS/MS	0.2 ± 0.2	0.2 ± 0.2
C12:1 (Dodecenoylcarnitine)	LC–MS/MS	0.12 ± 0.03	0.11 ± 0.02
C12 (Dodecanoylcarnitine)	LC–MS/MS	0.1 ± 0.1	0.04 ± 0.03
C14:2 (Tetradecadienylcarnitine)	LC–MS/MS	0.011 ± 0.002	0.009 ± 0.002
C14:1 (Tetradecenoylcarnitine)	LC–MS/MS	0.01 ± 0.01	0.01 ± 0.01
C14 (Tetradecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01	0.01 ± 0.01
C12-DC (Dodecanedioylcarnitine)	LC–MS/MS	0.017 ± 0.004	0.016 ± 0.002
C14:2-OH (Hydroxytetradecadienylcarnitine)	LC–MS/MS	0.006 ± 0.001	0.005 ± 0.001
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC–MS/MS	0.008 ± 0.004	0.009 ± 0.004
C16:2 (Hexadecadienylcarnitine)	LC–MS/MS	0.007 ± 0.001	0.007 ± 0.001
C16:1 (Hexadecenoylcarnitine)	LC–MS/MS	0.032 ± 0.004	0.032 ± 0.004
C16 (Hexadecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01	0.009 ± 0.004
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC–MS/MS	0.006 ± 0.001	0.007 ± 0.001
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC–MS/MS	0.011 ± 0.003	0.011 ± 0.004
C16-OH (Hydroxyhexadecanoylcarnitine)	LC–MS/MS	0.008 ± 0.002	0.007 ± 0.003
C18:2 (Octadecadienylcarnitine)	LC–MS/MS	0.004 ± 0.001	0.004 ± 0.001
C18:1 (Octadecenoylcarnitine)	LC–MS/MS	0.008 ± 0.004	0.007 ± 0.003
C18 (Octadecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01	0.006 ± 0.002
C18:1-OH (Hydroxyoctadecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002	0.01 ± 0.01
TRACE ELEMENTAL COMPOUNDS			
Lithium	ICP–MS	1.5 ± 0.4	1.5 ± 0.3
Boron	ICP–MS	<LOQ[1]	36 ± 18
Sodium	ICP–MS	10109 ± 2109	12037 ± 2852
Magnesium	ICP–MS	279 ± 84	286 ± 105
Phosphorus	ICP–MS	21915 ± 3094	21818 ± 2594
Potassium	ICP–MS	49300 ± 7605	51519 ± 7094
Titanium	ICP–MS	1.2 ± 0.2	1.3 ± 0.2
Vanadium	ICP–MS	0.013 ± 0.004	0.009 ± 0.001
Manganese	ICP–MS	0.12 ± 0.02	<LOQ
Iron	ICP–MS	10 ± 4	8 ± 3
Cobalt	ICP–MS	0.02 ± 0.01	0.022 ± 0.003
Copper	ICP–MS	0.25 ± 0.13	0.2 ± 0.1
Rubidium	ICP–MS	37 ± 7	40 ± 7
Strontium	ICP–MS	0.06 ± 0.01	<LOQ
Cesium	ICP–MS	0.03 ± 0.01	0.03 ± 0.01
Thallium	ICP–MS	0.0007 ± 0.0002	0.0009 ± 0.0001

[*] Compounds that exhibited good agreement with literature values; [1] LOQ: limit of quantification; [a] Jung et al., 2010 [201]; [b] Muroya et al., 2019 [202]; [c] Kodani et al., 2017 [203]; [d] Kim et al., 2016 [204].

Table 2. S3 List of liver metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g)

Metabolite	Platform	Concentration	Literature value
WATER-SOLUBLE COMPOUNDS			
AMINO ACIDS			
Alanine	LC–MS/MS & NMR	1388 ± 276	2366–3515 [a]
Arginine	LC–MS/MS	9 ± 5	
Asparagine	LC–MS/MS	169 ± 37	
Aspartate	LC–MS/MS & NMR	384 ± 165	
Beta-alanine	NMR	750 ± 137	
Citrulline	LC–MS/MS	20 ± 5	
Creatine	LC–MS/MS & NMR	1321 ± 464	
Glutamate	LC–MS/MS & NMR	4092 ± 928	6122–7999 [a]
Glutamine	LC–MS/MS & NMR	1433 ± 231	1911–2576 [a]
Glutathione	LC–MS/MS & NMR	892 ± 472	
Glycine	NMR	3201 ± 499	
Histidine	LC–MS/MS	382 ± 67	
Isoleucine	LC–MS/MS & NMR	178 ± 29	
Leucine	LC–MS/MS & NMR	426 ± 117	
Lysine	LC–MS/MS	115 ± 33	
Methionine	LC–MS/MS & NMR	54 ± 16	
Ornithine	LC–MS/MS & NMR	270 ± 63	984–1184 [a]
Phenylalanine	LC–MS/MS & NMR	138 ± 22	
Proline	LC–MS/MS	352 ± 67	
Serine	LC–MS/MS	422 ± 114	
Threonine	LC–MS/MS & NMR	377 ± 95	
Tryptophan	LC–MS/MS & NMR	44 ± 7	
Tyrosine	LC–MS/MS & NMR	84 ± 27	
Valine	LC–MS/MS & NMR	408 ± 73	
BIOGENIC AMINES			
Acetyl-ornithine	LC–MS/MS	23 ± 8	
Asymmetric-dimethylarginine	LC–MS/MS	1 ± 1	
Carnosine	LC–MS/MS & NMR	393 ± 130	
Creatinine	LC–MS/MS & NMR	58 ± 12	
Dopamine	LC–MS/MS	2 ± 2	
Histamine	LC–MS/MS	25 ± 16	
Kynurenine	LC–MS/MS	3 ± 2	
Methionine-sulfoxide	LC–MS/MS	13 ± 28	
Methylhistidine	LC–MS/MS	25 ± 5	
Putrescine	LC–MS/MS	0.2 ± 0.1	
Sarcosine	LC–MS/MS & NMR	23 ± 10	
Serotonin	LC–MS/MS	2 ± 1	
Taurine	LC–MS/MS & NMR	2243 ± 1186	
Total-dimethylarginine	LC–MS/MS	4 ± 1	

Trans-hydroxyproline	LC–MS/MS	51 ± 13
CARBOHYDRATES		
Glucose	LC–MS/MS & NMR	80098 ± 14629
ORGANIC ACIDS		
3-hydroxybutyrate	NMR	355 ± 89
Acetate	NMR	201 ± 91
Alpha-aminoadipate	LC–MS/MS	68 ± 27
Ascorbate (Vitamin C)	NMR	880 ± 257
Formate	NMR	1686 ± 209
Fumarate	NMR	299 ± 55
Inosinate (IMP)	NMR	34 ± 27
Lactate	NMR	12311 ± 1719
Nicotinurate	NMR	641 ± 87
Pyruvate	NMR	108 ± 20
Succinate	NMR	1194 ± 584
MISCELANEOUS		
Betaine	LC–MS/MS & NMR	358 ± 93
Choline	LC–MS/MS & NMR	345 ± 109
Ethanol	NMR	2065 ± 985
Glycerol	NMR	6219 ± 1151
Hypoxanthine	NMR	1227 ± 156
Inosine	NMR	833 ± 140
Myo-inositol	NMR	1177 ± 221
NAD+	NMR	347 ± 99
NADH	NMR	48 ± 26
NADP+	NMR	16 ± 8
O-phosphocholine	NMR	1658 ± 593
sn-Glycero-3-phosphocholine	NMR	10833 ± 1521
Uridine monophosphate (UMP)	NMR	58 ± 31
Uracil	NMR	76 ± 30
Uridine	NMR	344 ± 86
LIPID-LIKE COMPOUNDS		
PHOSPHATIDYLCHOLINES, ACYL-ALKYL		
PC ae (36:0)	LC–MS/MS	5 ± 2
PC ae (40:6)	LC–MS/MS	2 ± 1
PHOSPHATIDYLCHOLINES, DIACYL		
PC aa (32:2)	LC–MS/MS	3 ± 1
PC aa (36:6)	LC–MS/MS	3 ± 1
PC aa (36:0)	LC–MS/MS	17 ± 5
PC aa (38:6)	LC–MS/MS	6 ± 3
PC aa (38:0)	LC–MS/MS	1.6 ± 0.4
PC aa (40:6)	LC–MS/MS	14 ± 6
PC aa (40:2)	LC–MS/MS	0.7 ± 0.2
PC aa (40:1)	LC–MS/MS	0.6 ± 0.2

LYSOPHOSPHATIDYLCHOLINES, ACYL C		
LysoPC(14:0)	LC–MS/MS	0.09 ± 0.01
LysoPC(16:1)	LC–MS/MS	0.06 ± 0.02
LysoPC(16:0)	LC–MS/MS	0.7 ± 0.2
LysoPC(17:0)	LC–MS/MS	0.15 ± 0.03
LysoPC(18:2)	LC–MS/MS	1.1 ± 0.3
LysoPC(18:1)	LC–MS/MS	0.87 ± 0.23
LysoPC(18:0)	LC–MS/MS	2.1 ± 0.4
LysoPC(20:4)	LC–MS/MS	0.41 ± 0.11
LysoPC(20:3)	LC–MS/MS	0.132 ± 0.051
LysoPC(24:0)	LC–MS/MS	0.06 ± 0.01
LysoPC(26:1)	LC–MS/MS	0.04 ± 0.01
LysoPC(26:0)	LC–MS/MS	0.09 ± 0.03
LysoPC(28:1)	LC–MS/MS	0.1 ± 0.02
LysoPC(28:0)	LC–MS/MS	0.2 ± 0.1
SPHINGOMYELINS		
SM(16:1)	LC–MS/MS	0.9 ± 0.3
SM(16:0)	LC–MS/MS	22 ± 7
SM(18:1)	LC–MS/MS	3 ± 1
SM(18:0)	LC–MS/MS	6 ± 2
SM(20:2)	LC–MS/MS	0.8 ± 0.3
HYDROXYSPHINGOMYELINS		
SM(14:1(OH))	LC–MS/MS	2 ± 1
SM(16:1(OH))	LC–MS/MS	5 ± 1
SM(22:2(OH))	LC–MS/MS	4 ± 1
SM(22:1(OH))	LC–MS/MS	14 ± 3
SM(24:1(OH))	LC–MS/MS	4 ± 1
ACYLCARNITINES		
C0 (Carnitine)	LC–MS/MS	22 ± 6
C2 (Acetylcarnitine)	LC–MS/MS	5 ± 2
C3:1 (Propenoylcarnitine)	LC–MS/MS	0.2 ± 0.1
C3 (Propionylcarnitine)	LC–MS/MS	3 ± 1
C4:1 (Butenylcarnitine)	LC–MS/MS	0.08 ± 0.01
C4 (Butyrylcarnitine)	LC–MS/MS	0.36 ± 0.08
C3-OH (Hydroxypropionylcarnitine)	LC–MS/MS	0.07 ± 0.02
C5:1 (Tiglylcarnitine)	LC–MS/MS	0.022 ± 0.004
C5 (Valerylcarnitine)	LC–MS/MS	0.17 ± 0.04
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC–MS/MS	0.09 ± 0.02
C6:1 (Hexenoylcarnitine)	LC–MS/MS	0.17 ± 0.03
C6 (C4:1-DC) (Hexanoylcarnitine)	LC–MS/MS	0.3 ± 0.1
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine)	LC–MS/MS	0.09 ± 0.02
C5:1-DC (Glutaconylcarnitine)	LC–MS/MS	0.026 ± 0.004
C5-DC (C6-OH)(Glutarylcarnitine)	LC–MS/MS	0.2 ± 0.1
C8 (Octanoylcarnitine)	LC–MS/MS	0.09 ± 0.02

C5-M-DC (methylglutarylcarnitine)	LC–MS/MS	0.11 ± 0.02
C9 (Nonaylcarnitine)	LC–MS/MS	0.04 ± 0.01
C7-DC (Pimelylcarnitine)	LC–MS/MS	0.14 ± 0.11
C10:2 (Decadienylcarnitine)	LC–MS/MS	0.04 ± 0.01
C10:1 (Decenoylcarnitine)	LC–MS/MS	0.3 ± 0.1
C10 (Decanoylcarnitine)	LC–MS/MS	0.16 ± 0.03
C12:1 (Dodecenoylcarnitine)	LC–MS/MS	0.3 ± 0.1
C12 (Dodecanoylcarnitine)	LC–MS/MS	0.031 ± 0.004
C14:2 (Tetradecadienylcarnitine)	LC–MS/MS	0.011 ± 0.002
C14:1 (Tetradecenoylcarnitine)	LC–MS/MS	0.011 ± 0.002
C14 (Tetradecanoylcarnitine)	LC–MS/MS	0.011 ± 0.003
C12-DC (Dodecanedioylcarnitine)	LC–MS/MS	0.021 ± 0.003
C14:2-OH (Hydroxytetradecadienylcarnitine)	LC–MS/MS	0.008 ± 0.001
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002
C16:2 (Hexadecadienylcarnitine)	LC–MS/MS	0.008 ± 0.001
C16:1 (Hexadecenoylcarnitine)	LC–MS/MS	0.032 ± 0.003
C16 (Hexadecanoylcarnitine)	LC–MS/MS	0.012 ± 0.004
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC–MS/MS	0.02 ± 0.01
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC–MS/MS	0.021 ± 0.004
C16-OH (Hydroxyhexadecanoylcarnitine)	LC–MS/MS	0.011 ± 0.003
C18:2 (Octadecadienylcarnitine)	LC–MS/MS	0.006 ± 0.001
C18:1 (Octadecenoylcarnitine)	LC–MS/MS	0.011 ± 0.002
C18 (Octadecanoylcarnitine)	LC–MS/MS	0.008 ± 0.001
C18:1-OH (Hydroxyoctadecenoylcarnitine)	LC–MS/MS	0.011 ± 0.002
TRACE ELEMENTAL COMPOUNDS		
Lithium	ICP–MS	2 ± 1
Boron	ICP–MS	47 ± 28
Sodium	ICP–MS	25588 ± 4753
Magnesium	ICP–MS	10 ± 7
Phosphorus	ICP–MS	32820 ± 6825
Potassium	ICP–MS	46154 ± 8295
Calcium	ICP–MS	17 ± 8
Titanium	ICP–MS	2.2 ± 0.4
Cobalt	ICP–MS	0.8 ± 0.2
Copper	ICP–MS	27 ± 19
Zinc	ICP–MS	39 ± 60
Rubidium	ICP–MS	80 ± 16
Molybdenum	ICP–MS	3 ± 1
Cesium	ICP–MS	0.02 ± 0.01
Lead	ICP–MS	0.019 ± 0.003

[a] Miles et al., 2015 [205].

Table 2. S4 List of testis metabolites along with their measured concentrations and their standard deviations (as measured in nmol/g)

Metabolite	Platform	Concentration	Literature value
WATER-SOLUBLE COMPOUNDS			
AMINO ACIDS			
Alanine *	LC–MS/MS & NMR	1400 ± 294	880–1200 [a]
Arginine	LC–MS/MS	51 ± 12	200–460 [a]
Asparagine	LC–MS/MS	48 ± 11	
Aspartate *	LC–MS/MS & NMR	597 ± 215	310–550 [a]
Beta-alanine	NMR	48 ± 13	
Citrulline	LC–MS/MS	19 ± 7	
Creatine	LC–MS/MS & NMR	7553 ± 1850	
Glutamate	LC–MS/MS & NMR	3270 ± 702	1950–2430 [a]
Glutamine	LC–MS/MS & NMR	1517 ± 338	
Glutathione	NMR	1453 ± 257	
Glycine *	LC–MS/MS & NMR	1247 ± 251	930–1250 [a]
Histidine	LC–MS/MS	61 ± 13	20 [b]
Isoleucine	LC–MS/MS & NMR	79 ± 14	30 [b]
Leucine	LC–MS/MS & NMR	138 ± 23	40 [b]
Lysine *	LC–MS/MS	64 ± 15	40 [b], 120–200 [a]
Methionine	LC–MS/MS & NMR	26 ± 6	
Ornithine	LC–MS/MS	10 ± 3	
Phenylalanine	LC–MS/MS & NMR	56 ± 12	
Proline	LC–MS/MS	191 ± 40	
Serine *	LC–MS/MS	272 ± 58	230 [b], 630–830 [a]
Threonine	LC–MS/MS & NMR	232 ± 51	140 [b]
Tryptophan	LC–MS/MS & NMR	18 ± 4	
Tyrosine	LC–MS/MS & NMR	43 ± 9	
Valine	LC–MS/MS & NMR	197 ± 37	50 [b]
BIOGENIC AMINES			
Acetyl-ornithine	LC–MS/MS	7 ± 3	
Asymmetric-dimethylarginine	LC–MS/MS	0.8 ± 0.2	
Carnosine	LC–MS/MS	12 ± 3	
Creatinine	LC–MS/MS & NMR	123 ± 26	
Kynurenine	LC–MS/MS	0.9 ± 0.4	
Methionine-sulfoxide	LC–MS/MS	1 ± 0.4	
Methylhistidine	LC–MS/MS	19 ± 4	
O-phosphoethanolamine	NMR	6934 ± 1186	
Putrescine	LC–MS/MS	11 ± 3	
Sarcosine	LC–MS/MS & NMR	5 ± 2	
Spermidine	LC–MS/MS	0.4 ± 0.2	
Taurine	LC–MS/MS & NMR	1510 ± 296	
Total-dimethylarginine	LC–MS/MS	1.9 ± 0.4	
Trans-hydroxyproline	LC–MS/MS	43 ± 12	
Trimethylamine N-oxide	LC–MS/MS	11 ± 5	

CARBOHYDRATES		
Glucose	LC–MS/MS & NMR	149 ± 66
UDP-galactose	NMR	53 ± 10
UDP-glucose	NMR	126 ± 32
UDP-N-acetylglucosamine	NMR	268 ± 40
ORGANIC ACIDS		
3-hydroxybutyrate	NMR	101 ± 23
Acetate	NMR	88 ± 19
Alpha-aminoadipate	LC–MS/MS	98 ± 39
Ascorbate (Vitamin C)	NMR	1576 ± 363
Formate	NMR	823 ± 240
Fumarate	NMR	27 ± 8
Inosinate (IMP)	NMR	97 ± 19
Lactate	NMR	7702 ± 1649
Nicotinurate	NMR	38 ± 11
Pyruvate	NMR	11 ± 3
Succinate	NMR	340 ± 76
MISCELANEOUS		
Adenosine	NMR	150 ± 114
Betaine	LC–MS/MS & NMR	564 ± 134
Choline	LC–MS/MS & NMR	177 ± 71
Ethanol	NMR	125 ± 111
Glycerol	NMR	380 ± 112
Hypoxanthine	NMR	176 ± 66
Inosine	NMR	225 ± 95
Myo-inositol	NMR	7193 ± 1232
NAD+	NMR	151 ± 22
NADH	NMR	21 ± 8
NADP+	NMR	12 ± 3
O-phosphocholine	NMR	1058 ± 234
sn-Glycero-3-phosphocholine	NMR	702 ± 121
Uridine monophosphate (UMP)	NMR	56 ± 11
Uridine	NMR	156 ± 36
LIPID-LIKE COMPOUNDS		
PHOSPHATIDYLCHOLINES, ACYL-ALKYL		
PC ae (36:0)	LC–MS/MS	0.9 ± 0.2
PC ae (40:6)	LC–MS/MS	0.6 ± 0.1
PHOSPHATIDYLCHOLINES, DIACYL		
PC aa (32:2)	LC–MS/MS	0.5 ± 0.1
PC aa (36:6)	LC–MS/MS	0.4 ± 0.1
PC aa (36:0)	LC–MS/MS	15 ± 3
PC aa (38:6)	LC–MS/MS	21 ± 5
PC aa (38:0)	LC–MS/MS	0.5 ± 0.1
PC aa (40:6)	LC–MS/MS	2.7 ± 0.4
PC aa (40:2)	LC–MS/MS	0.12 ± 0.02

PC aa (40:1)	LC–MS/MS	0.07 ± 0.01
LYSOPHOSPHATIDYLCHOLINES, ACYL C		
LysoPC(14:0)	LC–MS/MS	0.09 ± 0.01
LysoPC(16:1)	LC–MS/MS	0.03 ± 0.01
LysoPC(16:0)	LC–MS/MS	0.7 ± 0.1
LysoPC(17:0)	LC–MS/MS	0.03 ± 0.01
LysoPC(18:2)	LC–MS/MS	0.41 ± 0.13
LysoPC(18:1)	LC–MS/MS	0.4 ± 0.1
LysoPC(18:0)	LC–MS/MS	0.16 ± 0.03
LysoPC(20:4)	LC–MS/MS	0.2 ± 0.1
LysoPC(20:3)	LC–MS/MS	0.04 ± 0.01
LysoPC(24:0)	LC–MS/MS	0.04 ± 0.01
LysoPC(26:1)	LC–MS/MS	0.04 ± 0.01
LysoPC(26:0)	LC–MS/MS	0.08 ± 0.01
LysoPC(28:1)	LC–MS/MS	0.06 ± 0.01
LysoPC(28:0)	LC–MS/MS	0.08 ± 0.01
SPHINGOMYELINS		
SM(16:1)	LC–MS/MS	0.6 ± 0.1
SM(16:0)	LC–MS/MS	42 ± 6
SM(18:1)	LC–MS/MS	0.5 ± 0.1
SM(18:0)	LC–MS/MS	5 ± 1
SM(20:2)	LC–MS/MS	0.3 ± 0.1
HYDROXYSPHINGOMYELINS		
SM(14:1(OH))	LC–MS/MS	1.3 ± 0.2
SM(16:1(OH))	LC–MS/MS	1.8 ± 0.3
SM(22:2(OH))	LC–MS/MS	1.1 ± 0.2
SM(22:1(OH))	LC–MS/MS	1.3 ± 0.3
SM(24:1(OH))	LC–MS/MS	0.25 ± 0.04
ACYLCARNITINES		
C0 (Carnitine)	LC–MS/MS	20 ± 6
C2 (Acetylcarnitine)	LC–MS/MS	40 ± 11
C3:1 (Propenoylcarnitine)	LC–MS/MS	0.77 ± 0.13
C3 (Propionylcarnitine)	LC–MS/MS	0.6 ± 0.2
C4:1 (Butenylcarnitine)	LC–MS/MS	0.02 ± 0.01
C4 (Butyrylcarnitine)	LC–MS/MS	2.8 ± 0.7
C3-OH (Hydroxypropionylcarnitine)	LC–MS/MS	0.06 ± 0.01
C5:1 (Tiglylcarnitine)	LC–MS/MS	0.021 ± 0.003
C5 (Valerylcarnitine)	LC–MS/MS	0.27 ± 0.09
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC–MS/MS	0.16 ± 0.04
C6:1 (Hexenoylcarnitine)	LC–MS/MS	0.03 ± 0.01
C6 (C4:1-DC) (Hexanoylcarnitine)	LC–MS/MS	0.21 ± 0.06
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine)	LC–MS/MS	0.08 ± 0.02
C5:1-DC (Glutaconylcarnitine)	LC–MS/MS	0.021 ± 0.004
C5-DC (C6-OH)(Glutarylcarnitine)	LC–MS/MS	0.021 ± 0.004
C8 (Octanoylcarnitine)	LC–MS/MS	0.06 ± 0.01

C5-M-DC (methylglutarylcarnitine)	LC–MS/MS	0.016 ± 0.003
C9 (Nonaylcarnitine)	LC–MS/MS	0.014 ± 0.002
C7-DC (Pimelylcarnitine)	LC–MS/MS	0.1 ± 0.1
C10:2 (Decadienylcarnitine)	LC–MS/MS	0.05 ± 0.01
C10:1 (Decenoylcarnitine)	LC–MS/MS	0.27 ± 0.04
C10 (Decanoylcarnitine)	LC–MS/MS	0.06 ± 0.01
C12:1 (Dodecenoylcarnitine)	LC–MS/MS	0.13 ± 0.01
C12 (Dodecanoylcarnitine)	LC–MS/MS	0.04 ± 0.01
C14:2 (Tetradecadienylcarnitine)	LC–MS/MS	0.011 ± 0.002
C14:1 (Tetradecenoylcarnitine)	LC–MS/MS	0.011 ± 0.003
C14 (Tetradecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01
C12-DC (Dodecanedioylcarnitine)	LC–MS/MS	0.013 ± 0.002
C14:2-OH (Hydroxytetradecadienylcarnitine)	LC–MS/MS	0.007 ± 0.001
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002
C16:2 (Hexadecadienylcarnitine)	LC–MS/MS	0.006 ± 0.001
C16:1 (Hexadecenoylcarnitine)	LC–MS/MS	0.032 ± 0.003
C16 (Hexadecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC–MS/MS	0.007 ± 0.001
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC–MS/MS	0.011 ± 0.001
C16-OH (Hydroxyhexadecanoylcarnitine)	LC–MS/MS	0.009 ± 0.001
C18:2 (Octadecadienylcarnitine)	LC–MS/MS	0.004 ± 0.001
C18:1 (Octadecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002
C18 (Octadecanoylcarnitine)	LC–MS/MS	0.02 ± 0.01
C18:1-OH (Hydroxyoctadecenoylcarnitine)	LC–MS/MS	0.009 ± 0.002
TRACE ELEMENTAL COMPOUNDS		
Lithium	ICP–MS	19 ± 2
Boron	ICP–MS	31 ± 11
Sodium	ICP–MS	35578 ± 6600
Magnesium	ICP–MS	70 ± 23
Phosphorus	ICP–MS	15602 ± 2618
Potassium	ICP–MS	48919 ± 8298
Calcium	ICP–MS	13 ± 4
Titanium	ICP–MS	1 ± 0.2
Cobalt	ICP–MS	0.05 ± 0.01
Copper	ICP–MS	0.7 ± 0.1
Zinc	ICP–MS	4 ± 1
Rubidium	ICP–MS	49 ± 9
Molybdenum	ICP–MS	0.06 ± 0.01
Cesium	ICP–MS	0.02 ± 0.01
Thallium	ICP–MS	0.0014 ± 0.0003
Lead	ICP–MS	0.014 ± 0.003

[*] Compounds that exhibited good agreement with literature values; [a] Brown-Woodman and White, 1974 [206]; [b] Sexton et al., 1971 [207].

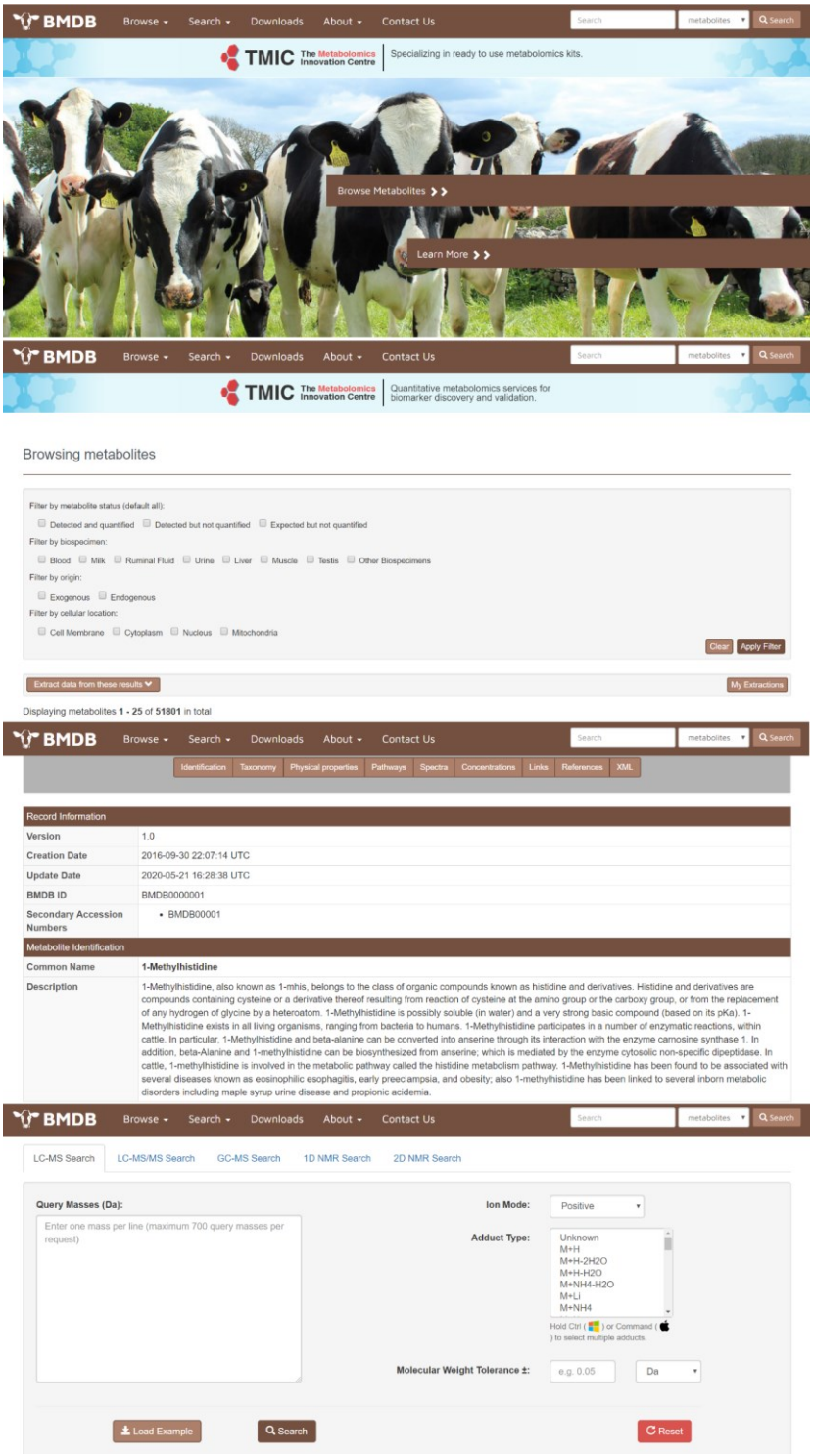

Figure 2. 1 Screenshot montage of different browsing and searching screens taken from the Bovine Metabolome Database (BMDB). A more detailed description of the different functions and capabilities of the various browse and search tools in the BMDB is given in the text.

CHAPTER 3. SERUM METABOLITE BIOMARKERS FOR PREDICTING RESIDUAL FEED INTAKE (RFI) OF YOUNG ANGUS BULLS

3.1. INTRODUCTION

Residual feed intake (RFI) is a livestock feed efficiency measure defined as the difference between an animal's actual feed intake and its expected feed requirements for maintenance and growth over a specific time-period. RFI is independent of growth characteristics such as body weight (BW) and average daily gain (ADG) [14, 16]. RFI measurements are laborious, expensive and time-consuming as they require measuring an individual animal's BW and feed intake over a period of 76 days [16, 33]. The RFI value is typically calculated over a group or herd of cattle, where the mean RFI for that group is defined as 0 kg/day. Low-RFI (LRFI) animals eat less than average, while high-RFI (HRFI) animals eat more than average. For example, an animal with an RFI value of -1.9 kg/day eats 1.9 kg/day less than the mean of 0 kg/day and is considered as a LRFI or a feed efficient animal. Selection for LRFI animals is gaining popularity among beef producers because LRFI cattle eat less per unit weight gain. Another positive attribute of LRFI cattle is that they produce less methane. Methane is a greenhouse gas, which is produced by ruminants during digestion and fermentation [208]. Livestock are responsible for the emission of ~18% of the global anthropogenic output of greenhouse gases [58, 208] and therefore reducing their carbon footprint is a key factor in reducing global warming. Several studies have shown that selecting for LRFI cattle is associated with reduced methane production [7, 17]. Indeed, compared to HRFI cattle, 25-28% lower methane production in LRFI animals has been reported [7, 17]. Therefore, selection for feed efficiency can favor both the farmer (decreased production costs) and the environment (lower methane and manure production). In addition, RFI has a moderate heritability (h^2 = 0.29–

0.46) in cattle, which makes it a good candidate for genetic improvement through selective breeding [16, 22, 28].

However, because RFI measurements are expensive and time-consuming, they are performed only on a small percentage of the cattle population. Simpler or cheaper proxies for measuring RFI are clearly desirable. Because RFI is a measure of metabolic efficiency, it has been proposed that metabolomics or metabolite measurements of bovine biofluids may offer a lower cost alternative to manual RFI measurement. Several metabolomics studies have been conducted in beef cattle to explore the relationship between RFI and metabolite levels [13, 23, 152]. For example, higher concentrations of glucose [23], urea [23], creatine [152], carnitine [152], and β-hydroxybutyrate [13], but lower concentrations of creatinine [23] were reported in the plasma of HRFI beef cattle compared to LRFI beef cattle. However, neither the performance of these biomarkers nor a precise mathematical model for predicting RFI from these biomarkers has been described. Likewise, these studies were limited to measuring a relatively small number of metabolites via a single metabolomics platform (such as nuclear magnetic resonance (NMR) spectroscopy) or a laboratory chemistry analyzer. Here we describe a more comprehensive metabolomic study that uses multiple metabolomics platforms, including NMR spectroscopy, liquid chromatography-tandem mass spectrometry (LC-MS/MS), and inductively coupled plasma-mass spectrometry (ICP-MS), to quantitatively characterize 145 serum metabolites in HRFI and LRFI young Angus bulls. Using this comprehensive metabolomics data set, we were able to identify several new metabolite biomarkers for RFI. Furthermore, we have constructed two logistic regression models (one optimized for NMR, the other optimized for LC-MS/MS) that use just two serum metabolites to differentiate HRFI and LRFI animals with a high sensitivity and specificity (AUROC > 0.85). A more complete description of the methods, the biomarkers and the models are given below.

3.2. MATERIALS AND METHODS

3.2.1. Ethics approvals

The collection and analysis of bovine serum in this study was approved by the University of Alberta's Animal Care Committee (Animal Use Protocol [AUP] 1129) under the auspices of the Canadian Council of Animal Care [209].

3.2.2. Animals and experimental design

Twenty-five purebred Angus bulls, raised on the University of Alberta's Roy Berg Kinsella Research Ranch (Kinsella, Alberta, Canada), were used in this study. After weaning, bulls were fed and managed according to industry standards for production of potential replacement yearling bulls in Alberta until their RFI test at approximately 13 months of age [73].

3.2.3. Measurement of phenotypic RFI values for the Angus bull cohort

From the end of May 2015 until mid-August 2015, bulls were tested for RFIf (RFI that was adjusted for rib fat thickness at the end of feedlot test) at approximately 13 to 16 months of age using the GrowSafe™ automated feed recording system (GrowSafe Systems Ltd.) at Agriculture and Agri-Food Canada (AAFC, Lacombe, Alberta, Canada). The RFI test was conducted following the protocols and calculation of RFI as reported by Mao et al. [210] and Johnson et al. [73], except for that standardized daily dry matter intake (STDDMI) was calculated as an average of dry matter intake over the test period and standardized to 12 megajoules of metabolizable energy (MJ ME) per kg dry matter for finishing bulls (instead of 10 MJ ME for heifers). The GrowSafe diet consisted of 45% barley and 55% silage (as fed basis), and the nutrient analysis is presented in Table 1. An adaptation period of 21 days was used to acclimatize cattle to the GrowSafe system and diet. The quantity of feed intake for each feeding event of each bull was recorded by the

GrowSafe system, which was further used to calculate total feed intake over the 77-day test period. Bulls were weighed twice at the beginning of test, once per month throughout the test, and once at slaughter, which was a few days after the RFI test was complete.

The end of RFI test weight was estimated from the slaughter weight. Rib fat thickness measurements (12/13th rib fat depth and LT area) were also determined at end of test, using an Aloka SSD-210 portable ultrasonographic scanner (Aloka Co., Tokyo, Japan). The initial BW at the start of the test and ADG were derived from a linear regression of the serial BW measurements against time (day). The metabolic BW (MWT) in kg was then calculated as midpoint $BW^{0.75}$, where the midpoint BW was computed as the sum of the initial BW and the product of ADG multiplied by half of the days on test. Using the dry matter (DM) content of the diet as well as the bull's daily intake, daily DMI in kg was calculated as an average of dry matter intake over the test period and was further standardized to 12 MJ ME per kg dry matter (STDDMI). In order to generate regression coefficients to predict an animal's expected DMI required for maintenance of body weight and growth, a linear regression model was fit using PROC GLM in SAS (SAS Institute, Inc., Cary, NC, USA). The model was:

$$Y_i = \beta_0 + \beta_1 ADG_i + \beta_2 MWT_j + \beta_3 FUFAT_k + e_{ijk} \{1\} \quad (1)$$

where Y_i is the STDDMI for the *i*th bull, β_0 is the intercept, β_1 is the partial linear regression coefficient of ADG, β2 is the partial linear regression coefficient of MWT, β_3 is the partial regression coefficient of final ultrasound backfat thickness (FUFAT), and e_i is residual error for the *i*th bull. RFIf in kg of DMI per day (kg/day) was computed as the difference between the standardized daily DMI and the expected DMI that was predicted based on animal's ADG, MWT and ultrasound backfat thickness in mm at the end of feedlot test (FUFAT) using the regression intercept and regression coefficients resulting from {1}. In total, 15 HRFI (0.39 ± 0.28 (mean RFIf

± standard deviation (SD))) and 10 LRFI (-0.52 ± 0.26) bulls were identified in this study. Those animals with RFIf value higher and lower than 0 kg/day were classified as HRFI and LRFI, respectively. RFIf values ranged from -1.05 kg to +1.07 kg DM per day, with an average of 0.00 kg/day.

3.2.4. Sample collection

Blood samples (10 mL) were collected in the morning (just before feeding) at 15 months of age from a jugular vein using vacutainer serum collection tubes (Becton Dickinson, Mississauga, Canada). Blood samples were kept in a cooler on ice, transferred to the laboratory within 3 hours after collection, and centrifuged at 2,000 ×g at 4 °C for 15 min. The upper layer of serum was then collected, and 4 mL was stored at -80 °C.

3.2.5. Metabolomics tests

Three metabolomics platforms, including NMR, LC-MS/MS and ICP-MS, were used to identify and quantify a total of 145 metabolites in each bovine serum sample. Using NMR, LC-MS/MS and ICP-MS, 42, 116, and 13 metabolites were identified and quantified respectively, of which 26 metabolites were common between NMR and LC-MS/MS. Details of sample preparation along with how the samples were run on each metabolomics platform have been previously described in detail by Foroutan et al. [211, 212]. Briefly, for NMR analysis, serum samples were filtered using a 3-kDa ultrafiltration unit (Amicon Micoron YM-3; Sigma-Aldrich, St. Louis, MO, USA) to remove large molecular weight proteins and lipoproteins. These macromolecules can seriously compromise the quality of ^{1}H-NMR spectra though the generation of intense, broad lines that interfere with the identification and quantification of lower abundance metabolites [213]. The de-proteinized sample was then frozen and stored at -80 ºC until further use. For NMR spectroscopic

analysis, 280 µL of the ultra-filtered serum was transferred to a 1.5 mL Eppendorf tube, to which an additional 70 µL of a standard NMR buffer solution (250 mM potassium phosphate (pH 7.0), 5 mM 2,2-dimethyl-2-silapentane-5 sulfonate (DSS-d_6), 5.84 mM 2-chloropyrimidine-5-carboxylic acid, and D_2O 54% v/v in H_2O) was added. The mixture was then transferred to a 3 mm NMR tube for spectral analysis. All ^{1}H-NMR spectra were collected on a Bruker Avance III Ascend 700 MHz spectrometer equipped with a 5 mm cryo-probe (Bruker Biospin, Rheinstetten, Germany). Compound identification and quantification by NMR were performed according the procedure described by Foroutan et al. [151], using the Chenomx NMR Suite 8.1 software package (Chenomx Inc., Edmonton, Canada).

A targeted, quantitative LC-MS/MS metabolite profiling approach was employed that combined reverse-phase liquid chromatography-mass spectrometry (RPLC-MS) with direct flow injection (DFI)-mass spectrometry (DFI-MS) (RPLC-DFI-MS/MS). LC-MS/MS was employed to determine the concentrations of up to 143 compounds (including amino acids, biogenic amines, glucose, organic acids, acylcarnitines, PCs, LysoPCs, SMs, and SM(OH)s) using an in-house quantitative metabolomics assay (TMIC Prime) [212, 214, 215]. All LC-MS analyses were conducted on an AB SCIEX QTRAP® 4000 mass spectrometer (Sciex Canada, Concord, Canada) equipped with an Agilent 1260 series UHPLC system (Agilent Technologies, Palo Alto, CA). The Analyst software 1.6.2 (Concord, Canada) was used to control the entire assay's workflow. The macro- and micro-elemental analyses were performed on a Perkin-Elmer NexION 350x ICP-MS (Perkin-Elmer, Woodbridge, Canada) according the procedure described by Foroutan et al. [212].

3.2.6. Statistical analysis

Data analysis was performed using MetaboAnalyst 4.0 according to previously published protocols [216, 217]. Those metabolites having more than two missing values in each group were

removed from further analyses. A univariate analysis including t-tests and fold-change analysis was performed in order to identify differentially expressed metabolites between the HRFI and LRFI groups. Statistical significance was declared at a p-value < 0.05.

Multivariate statistics, including PCA, PLS-DA, and ROC curve analysis, were performed using MetaboAnalyst 4.0. The data was scaled and normalized using a cube root transformation and auto scaling, which generated a clear Gaussian distribution plot prior to multivariate analysis. A permutation test involving 2000 randomized data sets was implemented to minimize the possibility that the observed separation of the PLS-DA was due to chance (a valid model should have a p-value < 0.05).

ROC curves were calculated by MetaboAnalyst 4.0 to evaluate the predictive ability of potential metabolic biomarkers using a logistic regression model. The area under the ROC curve (AUC or AUROC) was used to interpret the performance across the two different biomarker models to determine the best cut-off point for maximal sensitivity and specificity. A ROC curve plots the false-positive rate (1-specificity) on the X axis versus sensitivity on the Y axis. Sensitivity (or recall) is defined as the number of true positives divided by the sum of the true positives and false negatives. On the other hand, specificity is defined as the number of true negatives divided by the sum of the true negatives and false positives. In a ROC curve, the accuracy of a test for correctly distinguishing one group from another, such as HRFI bulls from LRFI bulls, is measured by the area under the ROC curve (AUROC). The AUROC equal to 1 is the highest value indicating a perfect discriminating test, which is obtained when all positive samples are ranked before negative ones. A permutation test involving 1000 randomized permutations was implemented to validate (a valid model should have a p-value < 0.05) the reliability of the model for each ROC curve.

3.3. RESULTS

3.3.1. The serum metabolome of beef cattle

Serum metabolomic data were obtained from 15 HRFI and 10 LRFI young Angus bulls using three metabolomics platforms including NMR, LC-MS/MS, and ICP-MS. A total of 145 metabolites were identified and quantified in each serum sample (Table 2). We have deposited this information into the Bovine Metabolome Database - BMDB (www.bovinedb.ca) [212]. Inspection of our experimental data reveals that the chemical composition of bovine serum is dominated by inorganic ions (primarily sodium, potassium, calcium, and phosphorus), carbohydrates (glucose), organic acids (lactate, acetate, and 3-hydroxybutyrate), amino acids (glycine, valine, and glutamine), and various amine-containing compounds (urea, creatinine). We found that for those metabolites that were measured by both LC-MS/MS and NMR, there was a generally good overall agreement with the concentration values across both platforms. Therefore, to simplify the presentation of the data we only report the LC-MS/MS values for those metabolites measured on both platforms. Based on our data, the range of metabolite concentrations detected in bovine serum varied from 1.2 ± 0.2 μM (fumarate) to 5393 ± 2341 μM (lactate) for NMR, from 0.0075 ± 0.0011 μM (C14:2-OH (hydroxytetradecadienylcarnitine)) to 4115 ± 326 μM (glucose) for LC-MS/MS, from 0.0016 ± 0.0001 μM (cesium) to 132919 ± 3122 μM (sodium) for ICP-MS.

Using a combination of NMR and LC-MS/MS, a total of 58 water-soluble organic compounds were identified and quantified in bovine serum. The most abundant water-soluble organic compounds in serum were lactate (5393 ± 2341 μM), glucose (4115 ± 326 μM), and urea (1389 ± 266 μM). The lowest concentration that could be reliably detected in serum was 0.035 ± 0.021 μM for putrescine.

The TMIC-Prime assay (a locally developed LC-MS/MS assay) provided quantitative results for 74 lipids or lipid-like compounds including 10 phosphatidylcholines (PCs), 14 lysophosphatidylcholines (LysoPCs), 5 sphingomyelins (SMs), 5 hydroxysphingomyelins (SM(OH)s), and 40 acylcarnitines (ACs) in bovine serum. Note that some LysoPC and PC species identified by the TMIC-Prime assay correspond to multiple (ranging from as few as 2 to as many as 24) possible unique lipid structures. In our study, SM(16:0) (69 ± 10 μM) and C14:2-OH (hydroxytetradecadienylcarnitine) (7.5 ± 1.1 nM) were the most and least abundant lipid-like compounds identified in serum, respectively.

ICP-MS also provided quantitative results for 13 trace minerals in bovine serum. The most abundant elements identified and quantified by ICP-MS were sodium (134 ± 16 mM), potassium (4.3 ± 0.3 mM), calcium (2.2 ± 0.2 mM), and phosphorus (1.3 ± 0.2 mM). While the least abundant metals quantified by ICP-MS were cesium (1.6 ± 0.2 nM), barium (190 ± 40 nM), and strontium (940 ± 140 nM).

3.3.2. Univariate statistical analysis of bovine serum metabolites

Using univariate analysis, we compared the serum metabolite profile of those young Angus bulls identified as being HRFI with those identified as being LRFI. The most significantly different metabolites (p-value < 0.05) between the HRFI and the LRFI animals are shown in Figure 1. In total, 10 differentially expressed metabolites achieved statistical significance in this comparison. Specifically, the serum concentrations of serine, leucine, formate, C0 (carnitine), C3 (propionylcarnitine), C4 (butyrylcarnitine), LysoPC(28:0), and SM(20:2) were greater in HRFI bulls than LRFI bulls. The most up-regulated metabolites were LysoPC(28:0) with a fold change (HRFI/LRFI) of 1.41 and C4 (butyrylcarnitine) with a fold change (HRFI/LRFI) of 1.38. In

addition to these eight up-regulated metabolites, two other metabolites, glycine and cesium, were down-regulated in the HRFI bulls as compared with their LRFI counterparts.

3.3.3. Multivariate analysis of bovine serum metabolites

Principle component analysis (PCA) showed moderately separable clustering between HRFI and LRFI animals (Figure 2a), while partial least squares-discriminant analysis (PLS-DA) showed a good separation for these two groups (Figure 2b). Permutation tests conducted on the PLS-DA model indicated that the observed separation was statistically significant (p-value < 0.01). A variable importance of projection (VIP) plot of the PLS-DA data, which ranks the top 15 metabolites based on their contribution to the discriminant model, is shown in Figure 3. The heat map on the right side of the VIP plot indicates that four metabolites (cesium, glycine, trimethylamine-N-oxide, and C10:2 (decadienylcarnitine)) were more abundant in the LRFI group, while the other 11 metabolites were more abundant in the HRFI group. All, except five metabolites (valine, trimethylamine-N-oxide, C10:2 (decadienylcarnitine), LysoPC(28:1), and acetyl-ornithine), identified via our multivariate analysis overlapped with the metabolites identified as significantly different between LRFI and HRFI animals by our univariate analysis.

3.3.4. Biomarkers for bovine RFI

From the significant metabolites identified via our univariate and multivariate analyses, we used logistic regression to generate two optimal models for distinguishing HRFI from LRFI animals. One biomarker panel uses only NMR-acquired data while the second uses only LC-MS/MS acquired data. The NMR model used two metabolites that are easily measured by NMR: formate and leucine (with an AUROC of 0.92 and a p-value of <0.01). The LC-MS/MS model also used

two metabolites that are easily measured by LC-MS/MS: C4 (butyrylcarnitine) and LysoPC(28:0) (with an AUROC of 0.89 and a p-value of <0.01).

As noted above, the best performing panel was the NMR-based test, which included formate and leucine. A logistic regression equation for these two candidate biomarkers was used to calculate the receiver operating characteristic curve (ROC) and to calculate the area under the ROC curve or AUROC (Figure 4). Permutation testing (n=1000) confirmed the significance of this model (p-value = 0.006). The logistic regression model developed for this prediction is given as follows:

$$\text{logit}(P) = \log(P / (1 - P)) = -1.6 - 3.554 \times \text{formate} - 2.161 \times \text{leucine} \quad (2)$$

where P is the probability of an animal being classified as LRFI. The optimal cutoff point for the above equation is 0.38. This means that an animal with a value greater than or equal to 0.38 belongs to the LRFI group, while an animal with a value less than 0.38 belongs to the HRFI group. Because the concentrations of the metabolites used in this study were cube-root transformed and then scaled via auto scaling, the value for formate in the above equation corresponds to the (cube root [formate]-4.2684)/0.1951 (where [formate] is the measured concentration of this compound in μM, as quantified by NMR). Likewise, the value for leucine corresponds to the (cube root [leucine]-5.9522)/0.2257 (where [leucine] is the measured concentration of this compound in μM, as quantified by NMR).

The second best performing RFI prediction panel included two metabolites that could only be measured by LC-MS/MS: C4 (butyrylcarnitine) and LysoPC(28:0). A logistic regression equation for these two candidate biomarkers was used to generate a model with a final AUROC of 0.89 (Figure 5). Permutation testing (n=1000) confirmed its significance (p-value = 0.005). The logistic regression model developed for this prediction is given as follows:

$$\text{logit}(P) = \log(P / (1 - P)) = -3.625 - 5.351\ \text{C4(butyrylcarnitine)} - 6.378\ \text{LysoPC(28:0)} \quad (3)$$

where P is the probability of an animal being classified as LRFI. The optimal cutoff point for the above equation is 0.53. This means that an animal with a value greater than or equal to 0.53 belongs to the LRFI group, while an animal with a value less than 0.53 belongs to the HRFI group. Because the concentrations of the metabolites used in this study were cube root transformed and then scaled via auto scaling method, the value for C4 (butyrylcarnitine) in the above equation corresponds to the (cube root [C4(butyrylcarnitine)]-0.5557)/0.0499 (where [C4(butyrylcarnitine)] is the measured concentration of this compound in µM, as quantified by LC-MS/MS). Likewise, the value for LysoPC(28:0) corresponds to the (cube root [LysoPC(28:0)]-0.6494)/0.0726 (where [LysoPC(28:0)] is the measured concentration of this compound in µM, as quantified by LC-MS/MS).

3.4. DISCUSSION

The main objective of this study was to identify candidate serum biomarker metabolites that could successfully discriminate HRFI cattle from LRFI cattle. To optimize the likelihood of identifying robust RFI biomarkers we used a combination of three quantitative metabolomics platforms (NMR, LC-MS/MS, and ICP-MS). Using these three platforms, we were able to identify and quantify a total of 145 metabolites, including 58 water-soluble organic compounds, 74 lipid-like compounds, as well as 13 metal ions. Overall, we found a very good agreement between the results of these 145 experimentally quantified metabolites with those of reported elsewhere (available in www.bovinedb.ca [212]). Indeed, the concentrations reported for serum in the BMDB agreed well with our experimental data. For instance, the value of asparagine reported by our study ranged from 21–30 µM, and for the literature-derived data it ranged from 20–33 µM. This widespread

agreement was not unexpected as serum/plasma must be highly stable and cannot vary much in its metabolite concentrations to ensure physiological homeostasis [218].

Of course, there were a few exceptions to this rule. The most variable metabolite reported in serum was betaine. The value of betaine reported by our study ranged from 131–205 μM, and the literature-reported values ranged from 14–26 μM [219]. This variation could be due to a number of factors, including differences in diet, sex, age, breed, sample work-up or extraction, sample storage protocols, analytical platforms and instrument sensitivity. We believe the most likely contributor to this difference is diet, as the amount of betaine in the diet of our beef cattle would be expected to be different than that of dairy cattle in the reference study of Artegoitia et al. [219]. Overall, there were very few outliers like betaine. Therefore, the good agreement for metabolite concentrations we obtained for the Angus bulls used in this study, with other cattle breeds suggests that the RFI biomarkers we discovered here should be transferrable to other breeds of beef cattle fed similar kinds of diets.

3.4.1. Comparison with literature-reported biomarkers of bovine RFI

To date there have been four other published metabolomic studies that have attempted to identify relationships between blood metabolite levels and bovine RFI [23, 152, 220, 221]. The study of Fitzsimons et al. [23] showed higher concentrations of glucose and urea and lower concentrations of creatinine in the plasma of HRFI vs. LRFI heifers. The study by Karisa et al. [152] reported higher concentrations of creatine, carnitine, formate, hydroxyisobutyrate, and tyrosine in the plasma of HRFI beef cattle along with higher concentrations of glycine in the plasma of LRFI beef cattle. Clemmons et al. [220] reported that the serum concentrations of pantothenate, homocysteine, glutamine, and carnitine were found to be associated with divergent RFI in beef steers, although no concentration values for these metabolites were reported in the Clemmons et

al. study. A very recent study conducted by Jorge-Smeding et al. [221] found that plasma metabolites that are directly (ornithine) or indirectly (aspartate, lysine, valine) associated with the urea cycle were correlated with RFI in Charolais heifers.

Table 3 summarizes these previous metabolomic findings and compares them with the findings reported here. As can be seen in this table, there was a good agreement between our findings and those reported from Fitzsimons et al. [23], Karisa et al. [152], and Jorge-Smeding [221]. For example, serum/plasma concentrations of tyrosine were higher in our HRFI group, which is in agreement with the findings of Fitzsimons et al. [23], and Karisa et al. [152]. Likewise, the serum concentrations of valine was higher in the HRFI group, which is similar to the findings of Fitzsimons et al. [23] and Jorge-Smeding et al. [221]. However, there were some discrepancies, with the most significant variations being seen in the study of Karisa et al. [152]. For instance, the serum concentration of 3-hydroxybutyrate was higher in HRFI animals in our study but reported as being lower is the study of Karisa et al. [152]. Karisa et al. [152] also reported exceptionally high concentrations for succinate (~250 μM), oxobutyrate (~40 μM), and allantonin (~90 μM), which do not match values reported by our study, by any other bovine studies or by the referential data in the BMDB [212]. Indeed, closer analysis of the NMR spectral regions corresponding to these metabolites (especially at higher fields) suggests that these peaks may have incorrectly identified and therefore incorrectly quantified. Other reasons for the differences between the Karisa et al. study and other bovine studies could be due to differences in diet, sex, age, breed, sample work-up or extraction or instrument sensitivity.

Another notable difference was found for blood glucose concentrations between our study and the values reported by Fitzsimons et al. [23]. In particular, the concentration of glucose was found to be higher in the serum of LRFI Angus bulls in our study but reported as being higher in the

plasma of Medium- and High-RFI Simmental heifers, respectively, in the study of Fitzsimons et al. [23]. Apart from glucose, other metabolites measured in both studies showed similar trends in terms of RFI classification (i.e. both studies found that the concentration of urea and carnitine were higher in HRFI animals). Glucose concentrations can vary significantly depending on how long samples are left at room temperature prior to being frozen. This is because glycolytic reactions in liquid serum/plasma readily lead to the conversion of glucose to lactate. Unfortunately, no details were provided in the study Fitzsimons et al. [23] regarding sample preparation time or lactate levels. Furthermore, given the fact that the highest concentration of glucose was seen in Medium-RFI animals as opposed to the LRFI or HRFI animals, suggests the glucose data reported by Fitzsimons et al. may have been more reflective of differences in sample preparation time than true differences in RFI. As a general rule, we treat reported glucose concentrations in livestock studies with a good deal of caution because of the extreme sensitivity of glucose levels to sample preparation/storage.

3.4.2. Candidate serum biomarkers of bovine RFI

While other studies have identified possible associations between blood metabolites and bovine RFI, as yet no published study has attempted to develop quantitative metabolite biomarker panels to predict RFI in cattle. Using logistic regression models, two categorical predictive biomarker panels were developed from this study to categorically predict RFI and to distinguish HRFI animals from LRFI animals.

The best performing panel was an NMR-based, two-metabolite model that included formate and leucine. The second best performing panel as an LC-MS/MS based two-metabolite model that included C4 (butyrylcarnitine) and LysoPC(28:0). Both panels have high sensitivity and specificity (AUROC > 0.85), making them good candidates to distinguish or predict HRFI animals from LRFI

animals. Because these panels consist of just two metabolites, it is possible to construct very fast (<5 minute/sample) and inexpensive (<$10) NMR or MS-based assays that could be used to perform bovine RFI characterization.

Basarab et al. [222] has estimated that a mature LRFI cow would have a net economic profit of $46/head/year compared to that of HRFI. This cost calculation suggests that selecting for LRFI could have a significant effect on reducing the costs of production for cattle ranchers. However, RFI is a difficult and time-consuming measure to perform. The cost of performing RFI measurements over 80-90 days is ~ $250/head which is much higher than the cost of a metabolite test ($5-10/head) or the net profit of selecting for LRFI cattle via GrowSafe™ RFI measurements. Therefore, a simple blood test that could distinguish high RFI (HRFI) animals from LRFI animals (early on) would potentially benefit beef farmers in terms of optimizing production or selecting which animals to cull or which animals should be bred.

As noted in the Methods section, the serum samples used to perform these metabolomic assays were collected at 15 months (shortly before the cattle were slaughtered at 17 months). Beef cattle produced in the United States and Canada can be slaughtered at any time from 12 months to 24 month of age, with the highest quality beef coming from those slaughtered under 24 months of age and the most tender meat found in animals slaughtered between 12-18 months of age [223]. Therefore, the markers identified here could be used for reasonably early prediction of RFI performance. However, it is not clear if the same panel of metabolites would work at other ages (14 months, 12 months, or 9 months) or whether the same panel would also work with cows, steers or heifers. Other metabolomic studies that have looked at metabolite-RFI associations [224, 225] suggest that these metabolic traits are likely established early in an animal's lifetime and so there is a good likelihood that these biomarkers could be used to assess RFI earlier than 15 months.

Being able to perform a serum-based RFI "prediction" even earlier in an animal's lifetime would certainly allow critical decisions to be made by the producer regarding breeding, culling or feeding a particular animal.

It is also important to note that other physiological factors certainly play a role in the composition of the bovine metabolome (and therefore the biomarker panel parameters described here), including physical maturity, sex and castration status. These physiological and age-dependent differences would be expected to lead to changes in the optimal cut-off concentrations. Typically, bulls are castrated at 3-6 weeks of age to become steers [226]. Uncastrated bulls reach puberty at 9-10 months [227] while heifers reach puberty at 12-14 months [228]. Castration will certainly affect some aspects of bovine metabolism as will the stage of an animal's sexual and physical maturity. Clearly additional studies will need to be done with other bovine cohorts over a range of ages and a range of physiological states to confirm the utility and cut-off values for these RFI biomarkers.

In addition to working with animals covering a wider set of ages (to ascertain the RFI biomarker age-range), it would also have been useful to perform further validation of these biomarkers on a "hold out" set of animals. These hold-out animals would have ideally raised elsewhere or at a different time using similar feeding, housing and animal management conditions. However, the high costs of measuring RFI, the length of the study (almost 2 years) and the costs of maintaining the animals for 17 months make these sorts of studies prohibitively expensive, especially given the limited resources for this sort of discovery-based study.

3.4.3. Metabolite markers and their role in RFI biochemistry

Our study identified a number of significantly different metabolites that seemed to drive the observed differences in RFI: C4 (butyrylcarnitine), LysoPC(28:0), formate, and leucine. Each of

these compounds plays an important role in bovine metabolism. C4 (butyrylcarnitine) is an acylcarnitine formed when fatty acyl-coenzyme A (fatty acyl-CoA) enters through the carnitine shuttle into the mitochondria for β-oxidation and the tricarboxylic acid (TCA) cycle to produce ATP [229]. Besides facilitating fatty acids crossing the mitochondrial membranes to be degraded by β-oxidation, acylcarnitines along with branched-chain amino acids or BCAAs (leucine, iso-leucine, and valine), also mediate activation of several important hepatic metabolic signaling pathways leading to diseases such as non-alcoholic fatty liver disease and type 2 diabetes mellitus in humans and other mammals [230, 231]. The short-chain acylcarnitines C3, C4, and C5 are degradation products of BCAAs [232] and saturation of the BCAA degradation pathway has been shown to inhibit the initial step of β-oxidation, leading to weight gain and body fat deposition [232]. High concentrations of BCAAs are associated with higher oxidative stress, and as seen in human and rodent studies, can serve as biomarkers for obesity-associated insulin resistance and diabetes [232, 233].

LysoPC(28:0) belongs to lysophosphatidylcholine family of lipids which are derived by partial hydrolysis of phosphatidylcholines by removing one of the fatty acid groups, via the action of phospholipase A2 (PLA2) [234]. High concentrations of lysoPC species (especially those containing palmitoyl (C16:0) or stearoyl (C18:0) groups) in the blood are known to stimulate cytosolic PLA2 and this results in an increased release of arachidonate, which is associated with cardiovascular disease [235]. In the vascular system, lysophosphatidylcholines have been shown to increase oxidative stress [236-238]. For example, Zou et al. [238] reported that lysophosphatidylcholines enhanced oxidative stress in rat aorta during aging via the 5-lipoxygenase pathway. Lehmann et al. [239] reported that circulating lysophosphatidylcholines can serve as biomarkers of a metabolically benign non-alcoholic fatty liver in humans. In

particular, Lehmann et al. found that the plasma concentration of lysophosphatidylcholines was higher in insulin-sensitive patients with non-alcoholic fatty liver compared with insulin-resistant ones with non-alcoholic fatty liver. Stiuso et al. [240] also reported that lipidomic and/or oxidative status of serum caused by lysophosphatidylcholines is associated with liver diseases (i.e. non-alcoholic fatty liver or steatohepatitis). A recent study [241] reported lower levels of reactive oxygen species (ROS) in the liver of LRFI steers which suggests they have lower levels of hepatic oxidative stress than HRFI steers. Decreased oxidative stress in the liver is associated with lower feed maintenance requirements, due to a lower lipid and protein turnover and better efficiency in energy usage [241].

As comprehensively reviewed by Herd and Arthur [19], variations in RFI can be explained by differences in energy expenditure from metabolic processes, body composition, and physical activity. Typically greater energy expenditures and higher maintenance requirements are seen in HRFI animals compared to LRFI animals [19]. Richardson et al. [30] also reported that Angus steers born from HRFI parents had less whole-body protein and more whole-body fat compared to progeny steers of LRFI parents. Therefore, higher levels of C4 (butyrylcarnitine) and LysoPC(28:0) in the serum of our HRFI bulls might be associated with increased oxidative stress in the HRFI group.

The other two serum metabolites that were most differentiating between HRFI and LRFI animals included formate and leucine. The concentration of formate and leucine was higher in HRFI animals and lower in LRFI animals. Formate participates in NADPH synthesis and catalyzes the conversion of fumarate into succinate in the TCA cycle [242, 243]. The associations between RFI and several metabolites (i.e., acetate, citrate and succinate) linked to the TCA cycle were recently discussed by Karisa et al. [152] as well as Wang and Kadarmideen [244]. Formate is the

simplest carboxylic acid and serves as a potent reductive force against oxidative stress. It is produced when the keto-acid, glyoxylate, neutralizes ROS in cells [242, 243]. Therefore, a higher level of formate in the serum of HRFI animals suggests that these less feed-efficient cattle are more prone to oxidative stress in the form of higher levels of ROS. This conclusion also agrees with the results reported in the study performed by Casal et al. [241]. Additionally, Fitzsimons et al. [23] also reported positive correlations between RFI and formate levels in the plasma of beef cattle.

Leucine is a branched-chain amino acid and its catabolism generates succinyl-CoA and acetyl-CoA, both of which can up-regulate the activity of the TCA cycle [245, 246]. BCAAs are also involved in protein turnover in skeletal muscle [246-248]. Leucine also increases fatty acid oxidation [249]. As discussed earlier, high concentrations of BCAAs are associated with higher levels of oxidative stress [232, 233]. Therefore, leucine could have an important role in RFI variation, since both protein turnover, oxidative stress and energy metabolism are key factors affecting this phenotype [19].

We also performed a further study to understand if the variations in the concentration of C4 (butyrylcarnitine), LysoPC(28:0), formate, and leucine between HRFI and LRFI bulls correlated with the concentration of these metabolites in their rumen. This was done to explore whether these metabolite difference may be associated with differences in ruminal activity or rumen microbial activity. However, we found no such correlation (data not shown).

3.5. CONCLUSIONS

In this study we evaluated the effectiveness of using multi-platform, quantitative metabolomics to identify candidate serum biomarkers that can easily distinguish HRFI animals from LRFI animals. LC-MS/MS, NMR and ICP-MS were used to identify and quantify 145 serum metabolites in an

effort to maximize our chances to identify and develop a suitable set of metabolite RFI biomarkers. We successfully identified two significant candidate biomarkers panels (AUROC > 0.85) that can predict RFI categorically. These include a two-metabolite model (formate and leucine) that is compatible with NMR analysis and a two-metabolite model (C4 (butyrylcarnitine) and LysoPC(28:0)) that is compatible with LC-MS/MS analysis. These results suggest that serum metabolites could be used to categorically predict RFI (early on) and inexpensively distinguish HRFI cattle from LRFI cattle.

While the results we obtained are very statistically significant and appear to be consistent with other reported studies on bovine RFI, the main limitation in this study was the small sample size (15 HRFI vs. 10 LRFI cattle). Given the significant costs and time associated with performing RFI measurements on cattle, this is a limitation that is difficult to overcome. Another limitation lies in the fact that the study was conducted on only a single sex (bulls), from a single breed (Angus cattle), consuming the same diet. However, it is important to note that we demonstrated that the data we measured in this study was broadly consistent with data collected for other beef cattle RFI studies. This gives us reason to believe that the results presented here will be shown to be largely reproducible elsewhere. Nevertheless, in order to properly confirm the robustness of these serum biomarkers as proxies to distinguish between divergent RFI cattle, further validation studies using a larger cohort of cattle with more diverse genetic backgrounds and from different management settings will be needed.

Table 3. 1 Nutrient analysis of barley-silage ration fed to bulls during the RFI test period in the GrowSafe system.

Diet composition	Value
DM [1]% (actual)	56.10
CP [2] (%DM)	14
ADF [3] (%DM)	25.25
NDF [4] (%DM)	40.50
TDN [5] (%DM)	69.60
Ca (%DM)	0.94
P (%DM)	0.34
Mg (%DM)	0.23
K (%DM)	1.38
Na (%DM)	0.13
Fe (PPM)	336
Mn (PPM)	70
Zn (PPM)	61
Cu (PPM)	16

[1] DM: Dry Matter Basis; [2] CP: Crude Protein; [3] ADF: Acid Detergent Fibre; [4] NDF: Neutral Detergent Fibre; [5] TDN: Total Digestible Nutrients

Table 3. 2 List of serum metabolites along with their analytical platform, measured concentrations, fold change, and Log2fold change.

Metabolite	Platform	HRFI (µM)	LRFI (µM)	Fold Change (HRFI/LRFI)	Log2Fold Change (HRFI/LRFI)
AMINO ACIDS					
Alanine	LC-MS/MS & NMR	236±29 [1]	245±30	0.96	-0.05
Arginine	LC-MS/MS & NMR	218±32	218±35	1.00	0.00
Asparagine	LC-MS/MS & NMR	26±4	24±3	1.08	0.12
Aspartate	LC-MS/MS & NMR	26±12	22±10	1.18	0.24
Beta-alanine	NMR	8±1	8±1	0.99	-0.02
Citrulline	LC-MS/MS & NMR	93±15	81±14	1.15	0.20
Creatine	LC-MS/MS & NMR	194±31	199±23	0.97	-0.04
Glutamate	LC-MS/MS & NMR	93±22	89±15	1.04	0.06
Glutamine	LC-MS/MS & NMR	330±52	330±22	1.00	0.00
Glycine *	LC-MS/MS & NMR	377±66	429±52	0.88	-0.19
Histidine	LC-MS/MS	78±12	79±8	0.99	-0.02
Isoleucine	LC-MS/MS & NMR	156±15	150±11	1.04	0.06
Leucine *	LC-MS/MS & NMR	221±25	197±15	1.12	0.17
Lysine	LC-MS/MS & NMR	91±18	84±9	1.08	0.12
Methionine	LC-MS/MS & NMR	33±5	34±3	0.97	-0.04
Ornithine	LC-MS/MS & NMR	60±13	63±12	0.95	-0.07
Phenylalanine	LC-MS/MS & NMR	72±7	69±7	1.04	0.06
Proline	LC-MS/MS & NMR	105±15	101±16	1.04	0.06
Serine *	LC-MS/MS & NMR	91±13	76±10	1.20	0.26
Threonine	LC-MS/MS & NMR	76±12	72±13	1.06	0.08
Tryptophan	LC-MS/MS	47±7	47±6	1.00	0.00
Tyrosine	LC-MS/MS & NMR	91±12	90±6	1.01	0.02
Valine	LC-MS/MS & NMR	367±33	338±28	1.09	0.12
BIOGENIC AMINES					
Acetyl-ornithine	LC-MS/MS	3.3±0.71	2.8±0.74	1.18	0.24
Asymmetric-dimethylarginine	LC-MS/MS	1.15±0.21	1.06±0.11	1.08	0.12
Carnosine	LC-MS/MS	31±16	29±6	1.07	0.10
Creatinine	LC-MS/MS & NMR	109±18	118±16	0.92	-0.11
Kynurenine	LC-MS/MS	7.3±1.2	7.6±2.4	0.96	-0.06
Methionine-sulfoxide	LC-MS/MS	1.2±0.3	1.2±0.3	1.00	0.00
Methylhistidine	LC-MS/MS	15±2	14±2	1.07	0.10
Putrescine	LC-MS/MS	0.035±0.021	0.041±0.014	0.85	-0.23
Sarcosine	LC-MS/MS & NMR	2.79±0.73	3.08±0.74	0.90	-0.15
Serotonin	LC-MS/MS	8±3	10±4	0.80	-0.32
Spermidine	LC-MS/MS	0.21±0.01	0.18±0.01	1.17	0.22
Spermine	LC-MS/MS	0.21±0.14	0.12±0.04	1.75	0.81
Taurine	LC-MS/MS & NMR	80±25	81±10	0.99	-0.02
Total-dimethylarginine	LC-MS/MS	2.1±0.3	2.1±0.3	1.00	0.00
Trans-hydroxyproline	LC-MS/MS	24±5	27±4	0.89	-0.17
Trimethylamine-N-oxide	LC-MS/MS	5±1	7±4	0.71	-0.49
CARBOHYDRATES					
Glucose	LC-MS/MS & NMR	3860±490	4115±326	0.94	-0.09
ORGANIC ACIDS					
3-hydroxybutyrate	NMR	375±164	287±94	1.31	0.39
Acetate	NMR	452±228	329±123	1.37	0.46
Alpha-aminoadipate	LC-MS/MS	1.25±0.54	1.31±0.44	0.95	-0.07
Ascorbate (Vitamin C)	NMR	11±3	10±3	1.10	0.14
Formate *	NMR	82±13	72±3	1.14	0.19
Fumarate	NMR	1.2±0.2	1.2±0.2	1.00	0.00
Lactate	NMR	4488±1761	5393±2341	0.83	-0.27
Pyruvate	NMR	142±27	162±54	0.88	-0.19
MISCELANEOUS					
Acetone	NMR	71±27	69±12	1.03	0.04
Betaine	LC-MS/MS & NMR	169±27	168±37	1.01	0.01
Choline	LC-MS/MS & NMR	20±4	22±4	0.91	-0.14
Ethanol	NMR	7.8±1.2	8.1±1.4	0.96	-0.05
Glycerol	NMR	312±41	318±36	0.98	-0.03
Isopropanol	NMR	2.27±0.82	2.54±0.34	0.92	-0.12
Methanol	NMR	32±5	31±3	1.03	0.05
Myo-inositol	NMR	43±12	48±8	0.90	-0.16
Urea	NMR	1389±266	1220±289	1.14	0.19
Uridine	NMR	3.1±0.71	2.8±0.52	1.11	0.15
PHOSPHATIDYLCHOLINES, ACYL-ALKYL					
PC ae (36:0)	LC-MS/MS	1.68±0.41	1.64±0.41	1.02	0.03

PC ae (40:6)	LC-MS/MS	0.47±0.13	0.44±0.04	1.07	0.10
PHOSPHATIDYLCHOLINES, DIACYL					
PC aa (32:2)	LC-MS/MS	4.3±1.3	3.8±1.1	1.13	0.18
PC aa (36:6)	LC-MS/MS	0.7±0.2	0.6±0.2	1.17	0.22
PC aa (36:0)	LC-MS/MS	6.05±1.4	6.18±1.4	0.98	-0.03
PC aa (38:6)	LC-MS/MS	1.007±0.284	0.901±0.194	1.12	0.16
PC aa (38:0)	LC-MS/MS	0.801±0.162	0.831±0.161	0.96	-0.05
PC aa (40:6)	LC-MS/MS	1.6±0.4	1.7±0.4	0.96	-0.05
PC aa (40:2)	LC-MS/MS	0.367±0.061	0.376±0.061	0.95	-0.08
PC aa (40:1)	LC-MS/MS	0.209±0.034	0.214±0.044	0.98	-0.03
LYSOPHOSPHATIDYLCHOLINES, ACYL C					
LysoPC(14:0)	LC-MS/MS	0.83±0.12	0.77±0.11	1.08	0.11
LysoPC(16:1)	LC-MS/MS	0.63±0.14	0.63±0.11	1.00	0.00
LysoPC(16:0)	LC-MS/MS	20±4	19±3	1.05	0.07
LysoPC(17:0)	LC-MS/MS	2.83±0.61	2.86±0.41	0.99	-0.02
LysoPC(18:2)	LC-MS/MS	16±4	14±2	1.14	0.19
LysoPC(18:1)	LC-MS/MS	6.5±1.4	6.4±1.1	1.02	0.02
LysoPC(18:0)	LC-MS/MS	29±6	30±3	0.97	-0.05
LysoPC(20:4)	LC-MS/MS	0.51±0.14	0.44±0.11	1.17	0.23
LysoPC(20:3)	LC-MS/MS	1.7±0.4	1.6±0.3	1.06	0.09
LysoPC(24:0)	LC-MS/MS	0.051±0.014	0.051±0.011	1.00	0.00
LysoPC(26:1)	LC-MS/MS	0.109±0.051	0.095±0.042	1.15	0.20
LysoPC(26:0)	LC-MS/MS	0.9±0.3	0.6±0.3	1.50	0.58
LysoPC(28:1)	LC-MS/MS	0.349±0.122	0.266±0.064	1.30	0.37
LysoPC(28:0) *	LC-MS/MS	0.322±0.121	0.228±0.044	1.41	0.50
SPHINGOMYELINS					
SM(16:1)	LC-MS/MS	6±1	5±1	1.10	0.13
SM(16:0)	LC-MS/MS	69±10	65±9	1.06	0.09
SM(18:1)	LC-MS/MS	11±3	9±2	1.22	0.29
SM(18:0)	LC-MS/MS	12±1	11±2	1.09	0.13
SM(20:2) *	LC-MS/MS	1.2±0.3	0.9±0.2	1.33	0.42
HYDROXYSPHINGOMYELINS					
SM(14:1(OH))	LC-MS/MS	5.6±1.2	5.1±1.1	1.10	0.13
SM(16:1(OH))	LC-MS/MS	9±1	8±2	1.13	0.17
SM(22:2(OH))	LC-MS/MS	5±1	4±1	1.10	0.13
SM(22:1(OH))	LC-MS/MS	9.3±1.4	8.8±1.4	1.06	0.08
SM(24:1(OH))	LC-MS/MS	1.9±0.4	1.9±0.4	1.00	0.00
ACYLCARNITINES					
C0 (Carnitine) *	LC-MS/MS	8±2	7±1	1.16	0.22
C2 (Acetylcarnitine)	LC-MS/MS	1.84±0.81	1.54±0.44	1.19	0.25
C3:1 (Propenoylcarnitine)	LC-MS/MS	0.028±0.004	0.029±0.004	0.97	-0.05
C3 (Propionylcarnitine) *	LC-MS/MS	0.22±0.052	0.18±0.014	1.22	0.29
C4:1 (Butenylcarnitine)	LC-MS/MS	0.017±0.002	0.017±0.002	1.00	0.00
C4 (Butyrylcarnitine) *	LC-MS/MS	0.197±0.041	0.143±0.034	1.38	0.46
C3-OH (Hydroxypropionylcarnitine)	LC-MS/MS	0.027±0.004	0.028±0.004	0.96	-0.05
C5:1 (Tiglylcarnitine)	LC-MS/MS	0.023±0.004	0.023±0.004	1.00	0.00
C5 (Valerylcarnitine)	LC-MS/MS	0.11±0.034	0.08±0.014	1.38	0.46
C4-OH (C3-DC) (Hydroxybutyrylcarnitine)	LC-MS/MS	0.041±0.004	0.042±0.004	0.98	-0.03
C6:1 (Hexenoylcarnitine)	LC-MS/MS	0.023±0.004	0.023±0.004	1.00	0.00
C6 (C4:1-DC) (Hexanoylcarnitine)	LC-MS/MS	0.053±0.014	0.049±0.011	1.08	0.11
C5-OH (C3-DC-M) (hydroxyvalerylcarnitine)	LC-MS/MS	0.038±0.004	0.036±0.004	1.06	0.08
C5:1-DC (Glutaconylcarnitine)	LC-MS/MS	0.018±0.003	0.018±0.003	1.00	0.00
C5-DC (C6-OH)(Glutarylcarnitine)	LC-MS/MS	0.028±0.004	0.027±0.004	1.04	0.05
C8 (Octanoylcarnitine)	LC-MS/MS	0.019±0.011	0.018±0.004	1.06	0.08
C5-M-DC (methylglutarylcarnitine)	LC-MS/MS	0.019±0.002	0.019±0.003	1.00	0.00
C9 (Nonaylcarnitine)	LC-MS/MS	0.022±0.002	0.021±0.003	1.05	0.07
C7-DC (Pimelylcarnitine)	LC-MS/MS	0.037±0.042	0.026±0.031	1.42	0.51
C10:2 (Decadienylcarnitine)	LC-MS/MS	0.05±0.01	0.06±0.01	0.89	-0.18
C10:1 (Decenoylcarnitine)	LC-MS/MS	0.172±0.032	0.163±0.034	1.06	0.08
C10 (Decanoylcarnitine)	LC-MS/MS	0.19±0.04	0.18±0.03	1.06	0.08
C12:1 (Dodecenoylcarnitine)	LC-MS/MS	0.085±0.013	0.081±0.014	1.05	0.07
C12 (Dodecanoylcarnitine)	LC-MS/MS	0.038±0.011	0.035±0.003	1.09	0.12
C14:2 (Tetradecadienylcarnitine)	LC-MS/MS	0.036±0.004	0.033±0.004	1.09	0.13
C14:1 (Tetradecenoylcarnitine)	LC-MS/MS	0.06±0.01	0.05±0.01	1.13	0.17
C14 (Tetradecanoylcarnitine)	LC-MS/MS	0.018±0.011	0.015±0.004	1.20	0.26
C12-DC (Dodecanedioylcarnitine)	LC-MS/MS	0.018±0.002	0.019±0.003	0.95	-0.08
C14:2-OH (Hydroxytetradecadienylcarnitine)	LC-MS/MS	0.0079±0.0021	0.0075±0.0011	1.05	0.07
C14:1-OH (Hydroxytetradecenoylcarnitine)	LC-MS/MS	0.008±0.002	0.009±0.001	0.89	-0.17
C16:2 (Hexadecadienylcarnitine)	LC-MS/MS	0.012±0.002	0.012±0.002	1.00	0.00
C16:1 (Hexadecenoylcarnitine)	LC-MS/MS	0.026±0.004	0.025±0.002	1.04	0.06

C16 (Hexadecanoylcarnitine)	LC-MS/MS	0.021±0.011	0.019±0.004	1.11	0.14
C16:2-OH (Hydroxyhexadecadienylcarnitine)	LC-MS/MS	0.005±0.001	0.006±0.001	0.83	-0.26
C16:1-OH (Hydroxyhexadecenoylcarnitine)	LC-MS/MS	0.018±0.003	0.019±0.004	0.95	-0.08
C16-OH (Hydroxyhexadecanoylcarnitine)	LC-MS/MS	0.007±0.001	0.008±0.001	0.88	-0.19
C18:2 (Octadecadienylcarnitine)	LC-MS/MS	0.006±0.001	0.007±0.001	0.86	-0.22
C18:1 (Octadecenoylcarnitine)	LC-MS/MS	0.014±0.003	0.016±0.003	0.88	-0.19
C18 (Octadecanoylcarnitine)	LC-MS/MS	0.022±0.011	0.021±0.004	1.10	0.14
C18:1-OH (Hydroxyoctadecenoylcarnitine)	LC-MS/MS	0.009±0.001	0.008±0.001	1.13	0.17
METAL IONS					
Sodium (Na)	ICP-MS	132919±12091	134408±16387	0.99	-0.02
Magnesium (Mg)	ICP-MS	920±77	948±104	0.97	-0.04
Phosphorus (P)	ICP-MS	1315±193	1271±111	1.03	0.05
Potassium (K)	ICP-MS	4283±428	4315±341	0.99	-0.01
Calcium (Ca)	ICP-MS	2251±232	2193±211	1.03	0.04
Iron (Fe)	ICP-MS	49±14	57±10	0.86	-0.22
Copper (Cu)	ICP-MS	8±2	9±2	0.89	-0.17
Zinc (Zn)	ICP-MS	13±2	12±1	1.05	0.07
Selenium (Se)	ICP-MS	1.4±0.2	1.3±0.2	1.08	0.11
Rubidium (Rb)	ICP-MS	1.8±0.2	1.8±0.2	1.00	0.00
Strontium (Sr)	ICP-MS	0.94±0.14	0.98±0.04	0.96	-0.06
Cesium (Cs) *	ICP-MS	0.0016±0.0002	0.0019±0.0003	0.84	-0.25
Barium (Ba)	ICP-MS	0.19±0.04	0.21±0.02	0.90	-0.14

[1] Mean ± Standard Deviation; * p-value < 0.05

Table 3. 3 Blood components associated with RFI as measured by different studies.

Metabolite	This study	Fitzsimons et al., 2013 (Olympus Chemistry Analyzer)	Karisa et al., 2014_Discovery population (NMR)	Karisa et al., 2014_Validation population (NMR)	Jorge-Smeding et al., 2019 (LC-MS/MS)
Glucose *	L [1]	H [2]			
Urea	H	H			
Creatinine	L	L	L		
Creatine *	L		H (Glutamine overlap)	H (Glutamine overlap)	
Carnitine	H	H	H (Glutamine overlap)	H (Glutamine overlap)	
Formate	H	H	L		
Hydroxyisobutyrate	ND [3]	H	H (Glucose overlap)		
Tyrosine	H	H	H		
Glycine	L	L	H		
Pantothenate	ND				
Hippurate	ND		H (Glutamine overlap)	L (Glutamine overlap)	
Threonine	H		H		
Acetate *	H		L		
Phenylalanine	H		H		
Lysine	H		H		
Citrate	ND		H		
Betaine	H		H		
Glutamate *	H		L		
Valine	H			H	H
Choline *	L			H	
Histidine	L			L	
Uridine	H			H	
2-methylamine	ND			L	
3-methylamine	ND			L	
2-hydroxybutyrate	ND			H	
3-hydroxybutyrate *	H			L	
4-hydroxybutyrate	ND			H (Acetone overlap)	
Succinate	ND			L (Mis-match)	
Oxo-butyrate	ND			L (Mis-match)	
Trans-4-hydroxy-L-proline	ND			L	
Proline	H			H	
Allantonin	ND			H (Mis-match)	
Glutamine	H=L			L (Overlap with glutamate, creatine, carnitine, hippurate)	
Aspartate	H				H
Ornitine *	L				H
Fumarate	H=L				L
Lysine	H				H

* Metobolites which concentration values do not agree with literature values; [1] L: LRFI; [2] H: HRFI; [3] ND: Not Detected

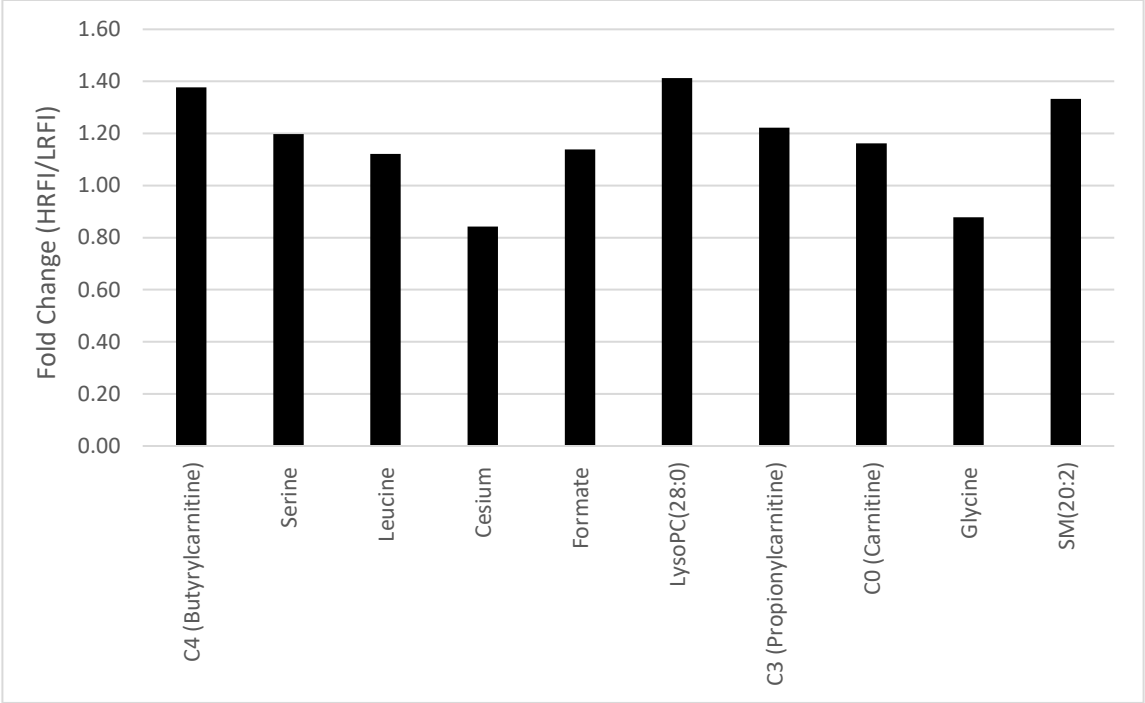

Figure 3. 1 Comparison of fold change of significantly regulated metabolites (p-value < 0.05) in the serum of HRFI versus LRFI bulls.

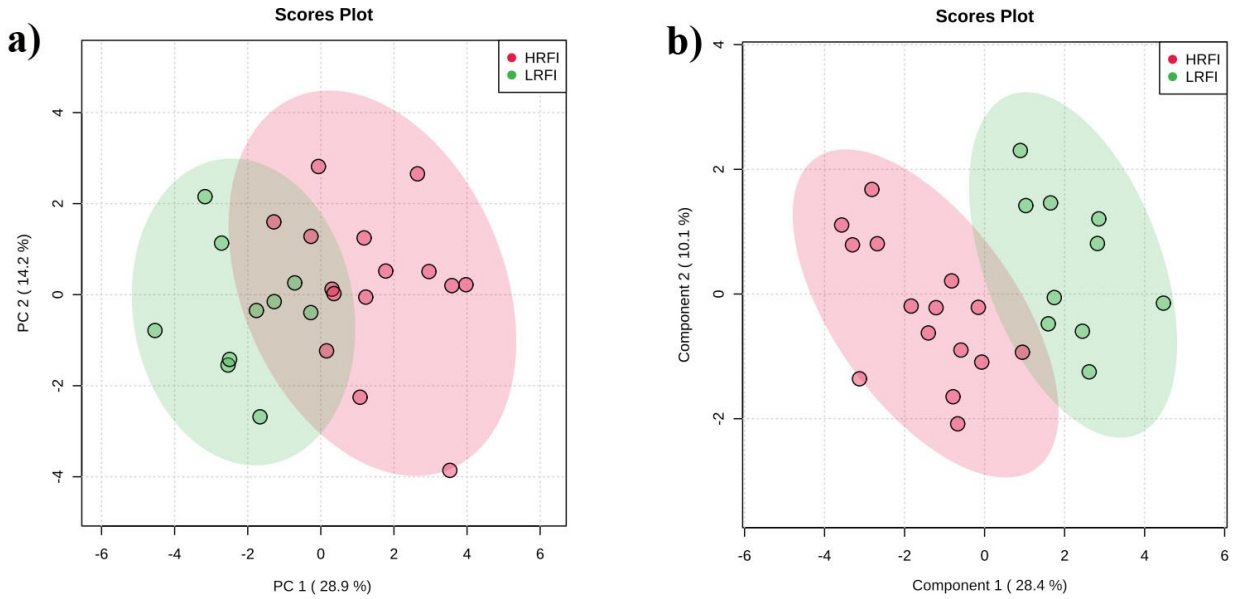

Figure 3. 2 Comparison between serum metabolite data acquired for HRFI versus LRFI group. **a**) Principal component analysis (PCA) graph. **b**) Partial least squares-discriminant analysis (PLS-DA) graph with permutation test p-value of < 0.01.

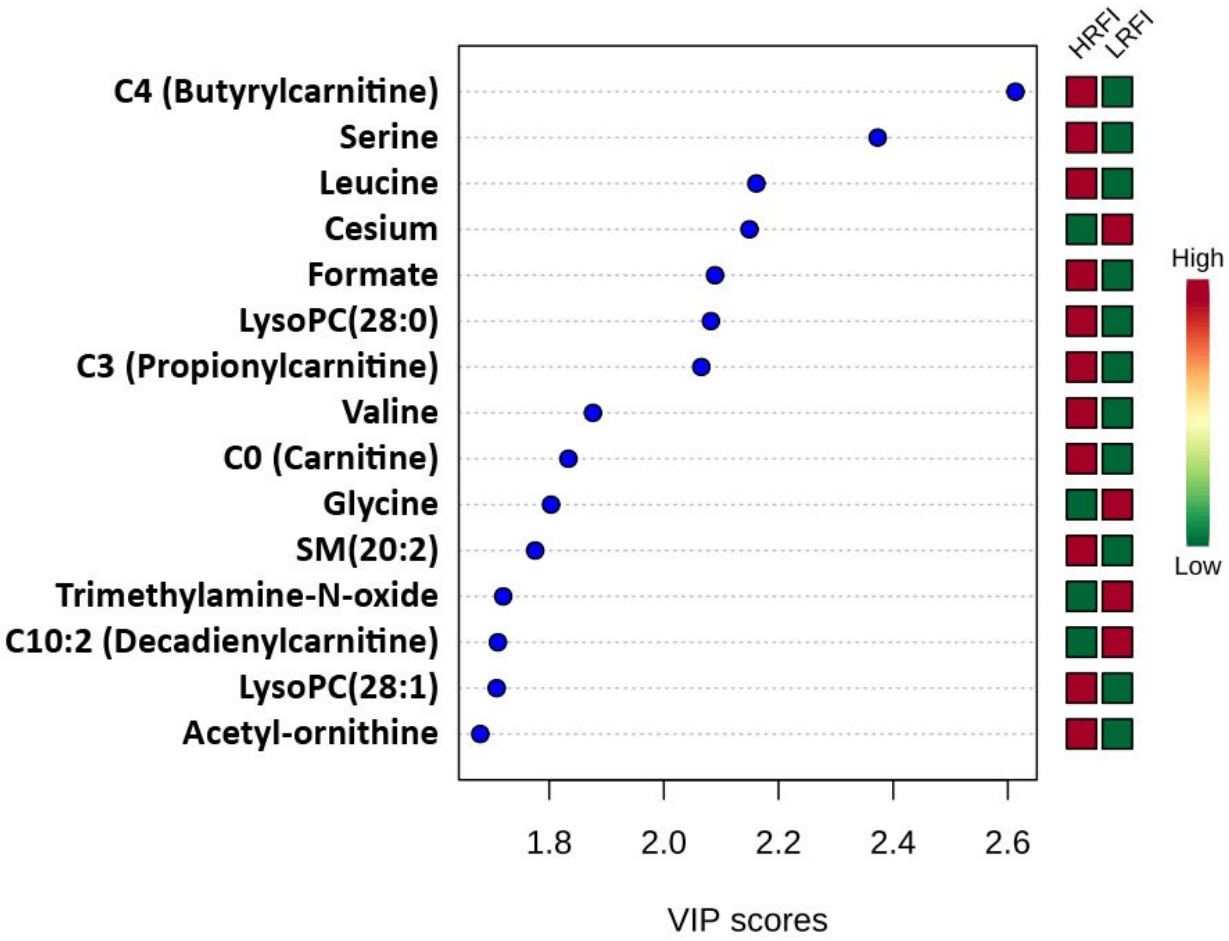

Figure 3. 3 Variable importance in projection (VIP) plot acquired from the comparison between HRFI vs. LRFI group. The most discriminating metabolites are shown in descending order of their coefficient scores. The color boxes indicate whether metabolite concentration is increased (red) or decreased (green) in HRFI vs. LRFI group.

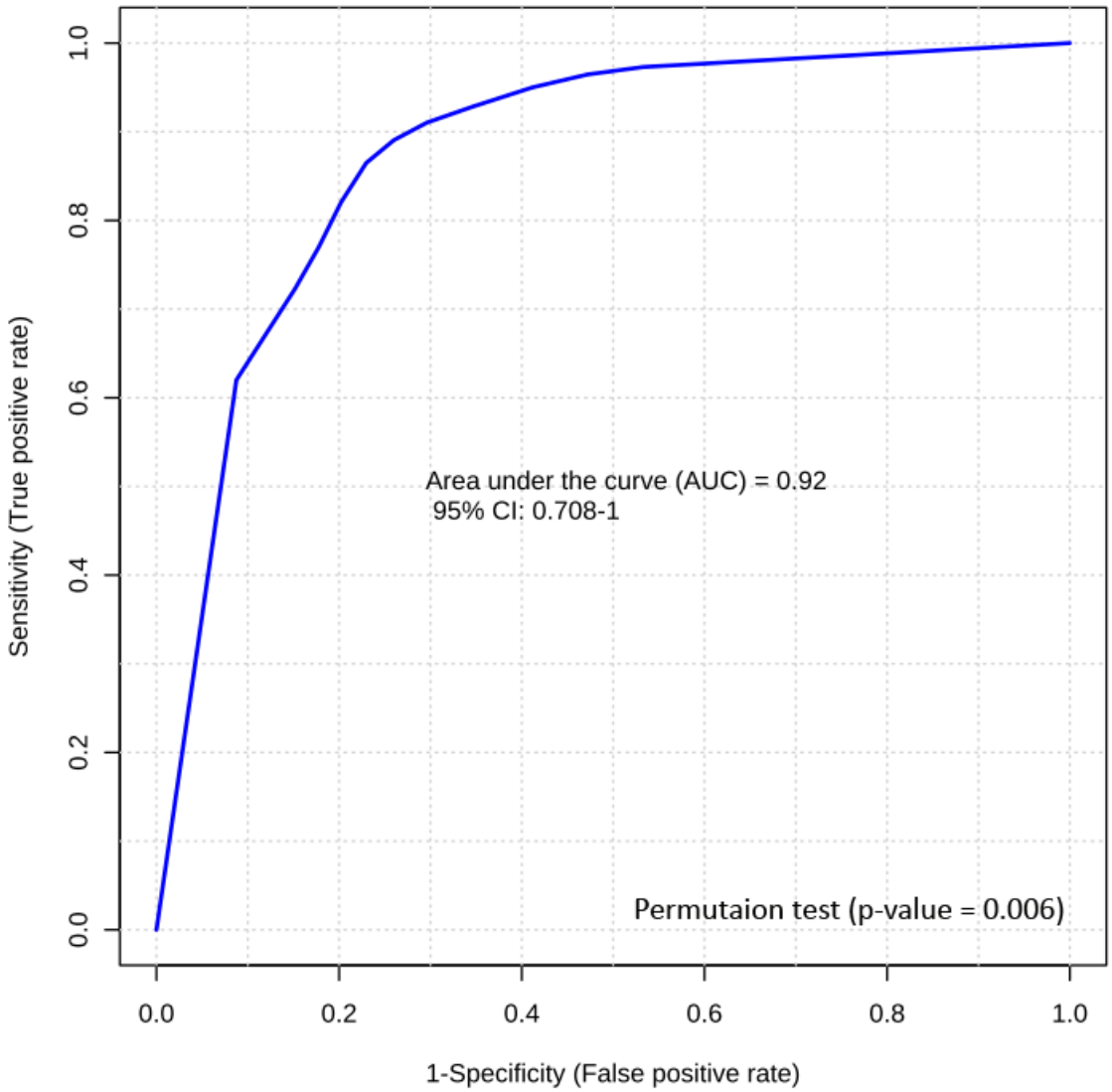

Figure 3. 4 Biomarker analysis of bovine RFI. Logistic regression ROC curve analysis of a panel of two NMR-detectable candidate biomarkers (formate and leucine) from bovine serum samples.

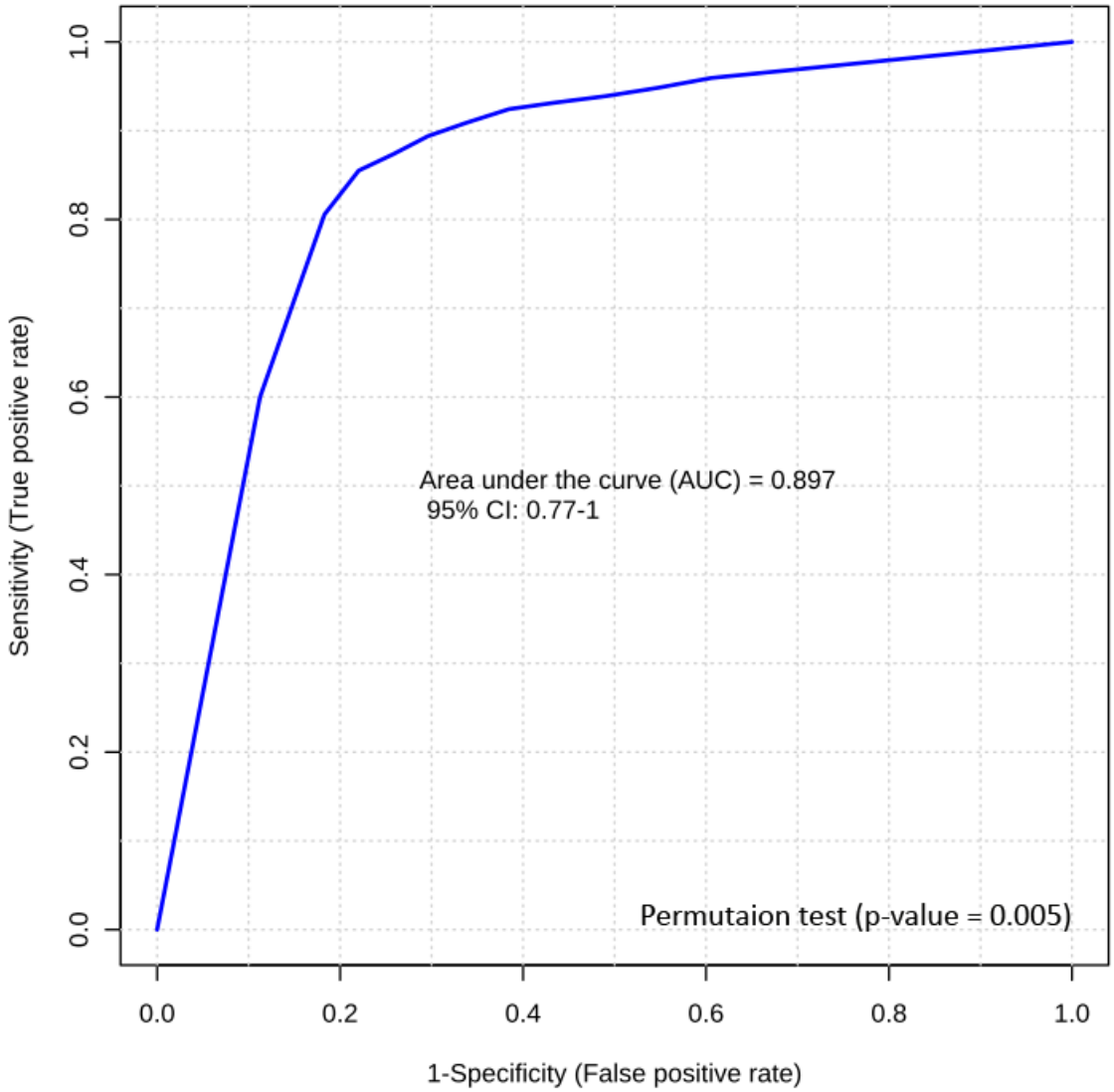

Figure 3. 5 Biomarker analysis of bovine RFI. Logistic regression ROC curve analysis of a panel of two LC-MS/MS candidate biomarkers (C4 (butyrylcarnitine) and LysoPC(28:0)) from bovine serum samples yields an AUC of 0.89.

CHAPTER 4. IMPACT OF PRENATAL MATERNAL NUTRITION AND PARENTAL RESIDUAL FEED INTAKE (RFI) ON MRNA ABUNDANCE OF METABOLIC DRIVERS OF GROWTH AND DEVELOPMENT IN YOUNG ANGUS BULLS

4.1. INTRODUCTION

In terms of the global competitiveness and sustainability of beef production, feed costs are among the biggest challenges, with ~70% of the cost of cow-calf production associated with winter feeding and pasture [250]. Consequently, decreasing feed costs could have a substantial benefit for the beef industry. There are several ways of decreasing feed costs in beef production, such as extending the grazing season and using alternative feeds [1], grouping cattle with similar nutritional needs together to optimize use of feedstuffs and reduce feed wastage [2], and analyzing feed composition to decrease expenses or generate additional profits through evidence-based decisions regarding buying or selling forages/grains [2]. Selection for feed efficient animals (Low-RFI or LRFI animals) is yet another strategy that not only reduces feed costs, but also decreases methane production per unit weight gain [7]. Therefore, selection for feed efficiency can both favor the farmer (decreased costs of production) and improve sustainability and environmental footprint of beef production.

Maternal malnutrition during gestation can affect postnatal growth and development, fertility, and the health of offspring [3-5, 85]. In beef cattle production, low to moderate undernutrition during the first half of gestation is considered of limited significance postnatally, since the fetus needs only limited nutrition for growth and development at this stage. However, critical events such as placental development and organogenesis occur during early pregnancy

[251]. In terms of nutrient partitioning during early gestation, the priority is with essential organs (i.e. heart, liver, lung, brain, kidney, etc.) rather than skeletal muscle [5, 252]. However, since no net increase in muscle fibre numbers occurs after birth, nutrient insults during early gestation may permanently alter an offspring's growth potential, and/or a variety of muscle characteristics important to beef production, including meat quality [3-5, 85, 95, 96]. However, male progeny lambs and calves born from dams that were nutritionally restricted during early- to mid-gestation, then re-alimented before parturition, had similar ratios of weight gain to feed intake compared to their respective control groups [5, 6, 253]. Hence, effects of substandard early prenatal nutrition in ruminants is not well established.

RFI is a feed efficiency measure defined as the difference between an animal's actual feed intake and its expected feed requirements for maintenance and growth over a specific time-period [14, 16]. Therefore, RFI is phenotypically independent of growth characteristics such as body weight (BW) and average daily gain (ADG). RFI has a moderate heritability (h^2 = 0.29–0.46) in cattle, making it a suitable candidate for genetic improvement [16, 22, 28]. Selection for LRFI animals is gaining popularity among beef producers as they eat less and produce less methane per unit weight gain compared to their High-RFI (HRFI) counterparts [7]. Basarab et al. [222] estimated that a mature LRFI cow would have a net economic profit of $46/head/year compared to that of an HRFI animal. However, it is important to investigate whether other traits, either positive or negative, might be co-selected along with LRFI. For example, if decreased reproductive potential is detected in LRFI cattle, selection for LFRI will require further economic evaluation. On the other hand, if LRFI is correlated positively with other beneficial traits, then selection for LRFI cattle will gain much more popularity and will clearly lead to increased profitability for the beef industry.

Studies that measure gene expression can reveal the biological mechanisms that are affected by prenatal diet, and that are important for variation in RFI. In beef cattle, gene expression differences in liver and muscle were associated with divergence in RFI [21, 160, 162, 254] and with prenatal nutrition [161, 194, 255, 256]. Although both RFI and maternal nutrition during gestation can affect progeny traits, our understanding of the biological mechanisms by which this occurs is poor. In particular, understanding of how RFI, maternal malnutrition during the first half of gestation, or their combination, can affect the production level (muscle development) and overall health of an animal is very important. In addition, we are uncertain as to how well LRFI cattle perform under differential environmental conditions relating to nutrient availability during gestation, among other factors. Therefore, this study aims to investigate the impact of selection for divergent genetic potential for RFI and maternal nutrition during early- to mid-gestation on beef cattle. More specifically, it explores how these two factors impact the expression of genes associated with metabolic drivers of growth and development and/or immunological function in the progeny Angus bulls.

4.2. MATERIAL AND METHODS

4.2.1. Ethics approvals

The collection and analysis of bovine tissues in this study were approved by the University of Alberta's Animal Care Committee (Animal Use Protocol (AUP) 1129) under the auspices of the Canadian Council of Animal Care [257].

4.2.2. Animals and experimental design

Details of heifer selection, experimental design, estrus synchronization and artificial insemination (AI), as well as prenatal dietary treatment, have been reported in detail by Johnson et al. [73], but

are described here. Eighty-nine purebred Angus heifers raised at the University of Alberta Roy Berg Kinsella Research Ranch (Kinsella, Alberta, Canada), were recruited for this study. The population in which these heifers originated were a herd of 240 purebred Angus cows maintained at the Kinsella Ranch with their pedigree information preserved by the Canadian Angus Association. All heifer calves had been vaccinated for clostridial diseases and viruses that are a part of the bovine viral disease complex. During late winter and early spring of 2013, heifers were tested for RFI at approximately 9 to 12 months of age over 74 days using the GrowSafe™ automated feed recording system (GrowSafe Systems Ltd., Airdrie, Alberta, Canada) as described by Basarab et al. [16] and Mao et al. [210], and is summarized in the following text. The test diet consisted of 70% silage and 30% barley grain, and a 21-day acclimatization period was allowed for heifers to adjust to the diet and GrowSafe feeding system before the intake measurements were officially recorded. The quantity of feed intake for each feeding event of each animal was recorded by the GrowSafe system, which was further used to calculate total feed intake over the test period. Feed samples were collected weekly and pooled monthly, and from the pooled samples dry matter (DM) and energy content analysis (Table 1) were determined using wet chemistry by Parkland laboratories (Red Deer, AB, Canada) in accordance with recommended methods of professional association of official Analytical Chemists, and the National Forage Testing Association. Heifers were weighed twice at the beginning and at the end of the 74-day test and at 14-day intervals throughout the test. Rib fat thickness measurements (12/13th rib and rib eye area) were determined at end of test, using an Aloka SSD-210 portable ultrasonographic scanner (Aloka Co., Tokyo, Japan). After the GrowSafe test, heifers were classified as either HRFI or LRFI by ordering the animals based upon their corresponding measurement of RFI, which was adjusted for rib fat thickness at the end of the feedlot test [37], from negative to positive, and then allocating them in

half. To produce calves with divergent potential for RFI, heifers classified as HRFI were bred to sires identified as HRFI (n = 2), whereas heifers classified as LRFI were bred to sires identified as LRFI (n = 2). Consequently, the progeny of these matings were expected to have divergent genetic potentials for RFI, due to the assortative breeding scheme. At the time of selection, the RFI estimated breeding values (EBVs) of the sires were +0.174 and +0.140 kg dry matter intake (DMI)/day (HRFI), and -0.230 and -0.482 kg DMI/day (LRFI), with accuracies of 0.859, 0.874, 0.805, and 0.712 respectively. Sires were also selected so that EBVs for other recorded traits were similar across all four bulls. Heifers within each RFI group were randomly stratified into two sub-groups, and each subgroup was assigned to one of the sires within their respective RFI classification (High versus Low) for breeding via estrus-synchronization and AI. Among the four heifer sub-groups, there were no biases in physical measurements collected during the RFI test such as end of test weight, ADG during test, metabolic mid-weight, or ultrasound backfat (data not shown). Also, there was no difference in standardized dry matter intake nor RFI for heifer groups assigned to AI sire within RFI classification (data not shown).

After the GrowSafe test, heifers were fed free-choice hay until mid-May, after which they grazed perennial mixed sown and native grass pasture during estrus-synchronisation and AI until pregnancy confirmation. Breeding was performed via two estrus-synchronized AI, as reported by Johnson et al. [73], but is also described here. The first AI was timed and the second was performed 12 hours after detection of estrus. On Day 0, a CIDR® (controlled internal drug release; Eazi-Breed^{TM}, Zoetis Animal Health, Kirkland, QC, Canada) was inserted intravaginally and 100 µg of gonadotropin-releasing hormone (GnRH; Fertiline®, Vétoquinol Canada Inc., Lavaltrie, QC, Canada) was given intramuscularly. On day 7, the CIDR® was removed and 500 µg of cloprostenol (Estrumate®, Merck Animal Health, Kirkland, QC, Canada) administered

intramuscularly. Fifty-five hours after CIDR removal and administration of Estrumate, a second injection of Fertiline® was administered intramuscularly and heifers were timed inseminated. To begin the second round of synchronisation, a CIDR® was again inserted into the vagina of each heifer on day 22 relative to the start of the 1st synchronisation, and remained there for 7 days. Neither Fertiline® nor Estrumate® were administered during the second round of estrous-synchronisation, as there was a possibility of heifers being pregnant. Concurrent with removal of the CIDR® on day 29, a KAMAR Heatmount™ detector (Kamar Inc., Steamboat Springs, CO, USA) was affixed with a liquid adhesive to aid in estrus detection. From Days 30 to 36, estrus detection was done three times daily via visual observation and heifers either observed in standing estrus and/or with an activated KAMAR were inseminated using the AM-PM rule (if heat first observed in the morning inseminated in the evening and vice versa).

Pregnancy diagnosis was performed at 28 days after each AI, via transrectal ultrasonography (Aloka-500V scanner equipped with a 7.5 MHz linear transducer; Aloka Co., Tokyo, Japan). Heifers that did not become pregnant after two rounds of AI were removed from the study. The remaining pregnant heifers (n = 61) were stratified by RFI, body weight, backfat depth, AI sire, and conception date. They were then randomly but equally assigned to one of two dietary treatments from d 30 to 150 of gestation (Table 1): 1) Low-diet (Ldiet) formulated for 0.5 kg/d ADG consisting of Brome grass hay (~75% of NRC requirements for growing pregnant heifers and of the normal diet), considered as a moderate level of restriction; or 2) Normal-diet (Ndiet) formulated for 0.7 kg/d ADG consisting of Brome grass hay supplemented with oats (roughly 100% of NRC requirements for growing, pregnant heifers). To account for heifer growth and increasing weight of the conceptus, the ration was adjusted approximately once per month after the heifers were weighed (Table 2). Heifers were fed the entire ration once a day, and if oats

were fed, it was provided in separate bunks before the hay was fed. Individual feed intake of Brome grass hay was recorded using the GrowSafe System. Supplemental salt and mineral were provided free-choice. Two heifers were removed from the prenatal nutrition trial after they were determined not to be pregnant after a subsequent pregnancy check.

After day 150 of gestation, all heifers were housed together and offered free choice hay (Table 1) until approximately 2 months after birth, when they grazed mixed tame and native grass pasture until weaning in November 2014 [73]. Male calves born to these heifers were left as bulls (8 Ndiet-HRFI, 9 Ldiet-HRFI, 3 Ndiet-LRFI, 5 Ldiet-LRFI). After weaning, bulls were fed and managed according to industry standards for production of potential replacement yearling bulls in Alberta until their RFI test at approximately 13 months of age [73].

4.2.3. Measurement of phenotypic RFI in bull progeny

From the end of May 2015 until mid-August 2015, bulls were tested for RFIf (RFI that was adjusted for rib fat thickness at the end of feedlot test) at approximately 13 to 16 months of age using the GrowSafe™ automated feed recording system (GrowSafe Systems Ltd.) at Agriculture and Agri-Food Canada (AAFC, Lacombe, Alberta, Canada). The RFI test was conducted as described for the heifers following the protocols and calculation of RFI as reported by Mao et al. [210] and Johnson et al. [73], except for that standardized daily dry matter intake (STDDMI) was calculated as an average of dry matter intake over the test period and standardized to 12 MJ ME per kg dry matter for finishing bulls (instead of 10 MJ ME for heifers). The GrowSafe diet consisted of 45% barley and 55% silage (as fed basis), and the nutrient analysis is presented in Table 3. An adaptation period of 21 days was used to acclimatize cattle to the GrowSafe system and diet. The quantity of feed intake for each feeding event of each bull was recorded by the GrowSafe system, which was further used to calculate total feed intake over the 77-day test period.

Bulls were weighed twice at the beginning of test, once per month throughout the test, and once at slaughter, which was a few days after the RFI test was complete. The end of RFI test weight was estimated from the slaughter weight. Rib fat thickness measurements (12/13th rib fat depth and rib eye area) were also determined at end of test, using an Aloka SSD-210 portable ultrasonographic scanner (Aloka Co., Tokyo, Japan). The initial BW at the start of the test and ADG were derived from a linear regression of the serial BW measurements against time (day). The metabolic BW (MWT) in kg was then calculated as midpoint $BW^{0.75}$, where the midpoint BW was computed as the sum of the initial BW and the product of ADG multiplied by half of the days on test. Using the dry matter (DM) content of the diet as well as the bull's daily intake, daily DMI in kg was calculated as an average of dry matter intake over the test period and was further standardized to 12 MJ ME per kg dry matter (STDDMI). In order to generate regression coefficients to predict an animal's expected DMI required for maintenance of body weight and growth, a linear regression model was fit using PROC GLM in SAS (SAS Institute, Inc., Cary, NC, USA). The model was:

$$Y_i = \beta_0 + \beta_1 ADG_i + \beta_2 MWT_j + \beta_3 FUFAT_k + e_{ijk} \quad \{1\}$$

Where Y_i is the STDDMI for the *i*th bull, β_0 is the intercept, β_1 is the partial linear regression coefficient of ADG, β2 is the partial linear regression coefficient of MWT, β_3 is the partial regression coefficient of final ultrasound backfat thickness (FUFAT), and e_i is residual error for the *i*th bull. After regression coefficients for ADG, MWT, and FUFAT were determined, expected DMI for each bull was determined by inserting the animal's own values into the equation {1} to obtain Y_i. After, RFIf in kg of DMI per day was computed as the difference between the actual standardized daily DMI as measured by GrowSafe, and the expected DMI (Y_i) that was predicted based on animal's ADG, MWT and FUFAT. RFIf was calculated as a difference between

the expected DMI and actual STDDMI. RFIf values of the bulls ranged from -1.05 kg to +1.07 kg DM per day, with an average of 0.00 kg.

After the RFI test and approximately about 17 months of age, the bulls were slaughtered at the AAFC-Lacombe Research Centre abattoir over four days, with one week separating the 2nd and 3rd slaughter dates [73]. On the day of slaughter, bulls were weighed to obtain their final weight (FW). At slaughter, they were first stunned by captive bolt and then exsanguinated. Following dressing and splitting of the carcass, hot side weights were recorded and added together to give the total trimmed weight (TW).

4.2.4. Tissue collection and RNA isolation

Between 5 to 10 g of tissue samples from LT muscle (from the left side of the bull, between the 12th and 13th ribs), SM muscle (left), liver, and testis were collected approximately 30-45 min after exsanguination. The tissue was immediately frozen in liquid nitrogen and then stored at -80°C until RNA isolation. Frozen tissue was ground by mortar and pestle in the presence of liquid nitrogen. Total RNA was isolated by the TRIzol procedure (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions with a few modifications. Fifty mg of ground tissue was transferred to a 2 mL metal, hard-tissue grinding beaded tube MK28 (Bertin Technologies, Montigny le Bretonneux, France), while the tube was on dry ice. Subsequently, 1 mL TRIzol reagent was added to the tube in order to perform homogenization and cell lysis. Tubes were then placed in the Cryolys® and Precellys® 24 tissue homogenization system (Bertin Technologies), and homogenization was performed twice at 5500 rpm for 30 s at 4 ºC, with an intervening 10 s pause. Tubes containing the homogenate were then centrifuged at 12,000 g for 10 min at 4 ºC, after which the supernatant was transferred via pipette to a new 2 ml Axygen™ MCT-200-O tube (Thermo Fisher Scientific, Waltham, MA, USA). Subsequently all RNA isolation steps were

completed according to the TRIzol RNA isolation procedure (Invitrogen). Finally, the RNA pellet was dissolved in 50 µL nuclease-free water (Ambion, Foster City, CA, USA) and then frozen at -80 ºC. After approximately 24 hrs, RNA concentration was quantified on a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), tested for integrity on an Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA), and stored at -80 °C until cDNA synthesis. All RNA samples used for RT-PCR had an RNA integrity number (RIN) score of ≥ 7.

4.2.5. Target gene selection

In order to explore how RFI and/or prenatal nutrition impact the expression of genes associated with growth and development (especially muscle development) as well as immunity in the progeny Angus bulls, 26 imprinted and non-imprinted target genes associated with these traits were selected from the literature (Table 4). As seen in Table 4, 12 genes associated with growth and development and/or immunity were tested in all four tissues. Muscle mass of livestock accounts for 35-60% of their body weight [258], and can serve as representative of overall growth. LT muscle mainly serves as a stabilizer rather than prime mover of the vertebral column as a whole, unlike SM muscle, which functions in locomotion and is highly active [258, 259]. This difference motivated us to compare how RFI and/or prenatal nutrition impacted the expression level of genes associated with growth and development between these two muscle tissues. We also surveyed the expression of genes associated with growth and development and/or immunity in liver since it is considered as the main centre of metabolism in the body, which breaks down various metabolites, synthesizes proteins, and produces biochemicals necessary for digestion and growth [260, 261]. In addition, we measured the expression of these 12 genes in testis since it expresses the highest number of genes compared to other tissues [262]. We hypothesized that expression of genes involved in growth and development and/or immunity in testis could be a reflection of baseline expression in

other tissues, with less of an influence of daily biological variation between individuals due to environmental impacts.

Apart from those 12 genes tested in all four tissues, an additional 14 were also tested only in liver. These 14 genes were expressed in the hepatic tissue of LRFI vs. HRFI cattle [162], and were involved in innate and adaptive immunity processes such as interferon signaling and inflammation. The liver is responsible for detoxifying and breaking down harmful chemicals. This contributes to its active role in immune function and failure to do this function can result in a higher concentration of circulating endotoxins, causing local and systemic inflammation [260, 261]. This has been shown to antagonize livestock's growth and performance since nutrients are partitioned for immunity processes rather than anabolic processes such as muscle synthesis [263, 264]. Therefore, we decided to check the expression of these 14 genes in the liver of our animals.

4.2.6. Designing and testing of real time-PCR primers

Primers were designed using Primer3 software (Version 0.4.0, http://bioinfo.ut.ee/primer3-0.4.0/) [265, 266], using species-specific sequences found in GenBank and were designed to cover exon-exon junctions when possible (Table 4). An oligo analyzer tool (Integrated DNA Technologies, https://www.idtdna.com/) was used to select primers with minimum probability of having hairpins, self-dimers, and hetero-dimers. For all RT-PCR reactions, the annealing/extension temperature was 60 ºC. The amplification efficiency for each gene was determined using serial dilution of tissue-specific cDNA and was 100 ± 10% for all genes (data not shown). For all the primer-pairs used in the final analyses, the product identity was confirmed by sequencing PCR products (data not shown).

For each tissue, seven endogenous genes including ribosomal protein L19 [*RPL19*], beta actin [*ACTB*], peptidylprolyl isomerase A (cyclophilin A) [*PPIA*], eukaryotic translation

elongation factor 1 alpha 2 [*EEF1A2*], glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], hydroxymethylbilane synthase [*HMBS*], tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta [*YWHAZ*] (Table 4) were tested to determine the best gene or combination of genes (the geometric mean of the combination of genes) using NormFinder software, version 20 [267]. *EEF1A2*, *HMBS*, the geometric mean of *GAPDH* and *HMBS* and *YWHAZ* and *PPIA* were selected as the best endogenous genes to normalize data for LT, SM, liver, and testis tissues, respectively. These best endogenous genes were also tested for any treatment effect and were determined to be stable amongst all samples within each tissue type confirming their suitability as endogenous controls.

4.2.7. cDNA creation and real time-PCR

cDNA synthesis was performed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions, with 1.5 µg of total RNA used for each individual sample or reference RNA. Each tissue's reference RNA was created by pooling 1 µL of every individual RNA sample (concentration 150 ng/µL) for that tissue.

For each of the four tissues, a standard curve was created by serial dilution of the tissue-specific cDNA reference sample, with reference to the original input RNA concentration (5, 1, 0.2, 0.04, 0.008, and 0.0016 ng input RNA). The serial dilution for the standard curve was made fresh from aliquoted reference cDNA before each RT-PCR reaction. Individual sample cDNAs for each tissue were diluted to a final concentration equal to 0.5 ng input RNA for the RT-PCR reaction. RT-PCR was carried out in a 96-well optical reaction plate (Applied Biosystems) using the Applied Biosystems Step-One-Plus RT-PCR system (Applied Biosystems). The total volume of each reaction was 20 µL and contained 4 µL of cDNA template (1 ng/µL), 0.5 µL of each forward

and reverse primer (10 μM), 10 μL of the 2× KAPA SYBR FAST qPCR Master Mix ABI Prism (Kapa Biosystems, Boston, MA, USA), and 5 μL of nuclease-free water (Ambion). Control samples in each run consisted of the reference RNA without conversion to cDNA by reverse transcriptase (no RT control), and water instead of template cDNA (no template control). All samples, standard curves, and controls were run in duplicate on each plate. Amplification conditions consisted of: (1) hold stage for 30 s at 95 °C for enzyme activation; (2) cDNA amplification stage for 40 cycles of 3 s at 95 °C and then 30 s at 60 °C; and (3) a final melt curve stage which consisted of 15 s at 95 °C, 1 min at 60°C. The PCR program finished with a melt of the reaction with a ramp rate of +0.3 °C per 15 s from 60 °C to 95 °C to obtain fluorescence measurements that reflect the T_m of the main product while minimizing interference from primer-dimer or other non-specific amplification products.

4.2.8. Real time-PCR data and statistical analysis

Ct values were converted to relative input RNA concentrations using the reaction efficiency of the standard curve in each run. For all genes (both target and endogenous), sample duplicates were averaged and expressed relative to the highest expression sample. These relative values were divided by the appropriate normalization gene value for that sample. Normalized gene expression values were analyzed as a general linear model (GLM) within SAS (Version 9.4). The model included parental RFI group (high and low), diet group (normal and low), and parental RFI×diet interaction as fixed independent variables. Bull age at slaughter was tested as a covariate, which was subsequently removed, as its contribution was not significant. If normalized gene expression values for individual genes/tissue were not normally distributed, the Box-Cox transformation within PROC TRANSREG was used to find the most appropriate data transformation. Differences between means were analyzed using a least significant difference (LSD) test with a 95%

confidence level and were reported as least squares means (LSM) ± the standard error of the mean (SEM). All LSM ± SEM reported are calculated from non-transformed data. The p-value for significance was < 0.05. Dependencies between individual gene expression within each tissue were evaluated by Pearson Correlation coefficients using the PROC CORR procedure in SAS (Version 9.4) to assess the need for multiple testing correction. Only genes that achieved acceptable amplification, including agreement between duplicates, a single peak in the melt curve, no amplification for control samples, plus a reaction efficiency between 90-110%, were analysed statistically.

4.2.9. Correlations between phenotypic and gene expression data

Pearson correlation analysis was performed between gene expression data in liver, testis, LT, and SM muscles, and phenotypic data of the progeny bulls including RFIf (kg/d), STDDMI (kg/d), ADG (kg/d), MWT (kg), FW (kg), and TW (kg), which were collected during the RFI test and at slaughter. Gene expression data that were not normal were transformed using the Box-Cox transformation within PROC TRANSREG, as appropriate. Pearson Correlation coefficients were calculated using the PROC CORR procedure in SAS (Version 9.4) to analyze the association between gene expression and the phenotypic data. STDDMI, MWT, FW, and TW were adjusted for age within PROC CORR as these measurements were significantly associated with age. P-values of < 0.05 were considered significant.

4.3. RESULTS

RT-PCR identified two DE genes in liver. Prenatal maternal diet was associated with expression of protocadherin 19 [*PCDH19*] (P = 0.01, Figure 1), where bull progeny of the Ldiet displayed higher expression than the Ndiet group. The LRFI group was associated with the expression of

myocyte enhancer factor 2A [*MEF2A*] (P = 0.01, Figure 2), where bull progeny of the LRFI displayed higher expression than the HRFI group. None of the other genes tested in liver were significantly different with respect to RFI or parental diet groups, nor their interaction (Table 5). Although we did see differences in expression of some genes that could be considered trends (P < 0.1) (Table 5), the p-values were very close to 0.1. Several significant correlations between expression of individual genes within each tissue were detected at P < 0.05 (27, 44, 78, and 50% of gene pair-wise correlation coefficients in liver, LT muscle, SM muscle, and testis, respectively, correlation data not shown). Therefore, P-value correction for multiple testing was not performed when detecting significant gene expression as it would result in a higher number of false negatives. Nevertheless, the identified postnatal gene expression responses to prenatal nutrition would require further validation in different beef cattle populations.

Correlation analysis between gene expression in liver and phenotypic traits recorded revealed several significant associations. STDDMI and MWT showed positive relationships with *PCDH19* (r = +0.56, P = 0.004; and r = +0.47, P = 0.019, respectively). ADG was negatively correlated with insulin like growth factor 1 receptor [*IGF1R*] (r = -0.44, P = 0.022). As well, a significant relationship between ADG and H19 imprinted maternally expressed transcript [*H19*] (r = +0.59 (transformed *H19*), P = 0.001) was observed, and a positive relationship was found between RFIf and STDDMI with S-transferase mu 1 and mu 2 [*GSTM1/2*] (RFIf: r = +0.51, P = 0.007; and STDDMI r = +0.42, P = 0.035). Furthermore, MWT was correlated with insulin like growth factor 2 receptor [*IGF2R*] (r = -0.41 (transformed *IGF2R*), P = 0.042).

Regarding LT and SM muscles, RT-PCR identified *MEF2A* as differentially expressed. In LT muscle, a maternal prenatal diet×RFI parental group interaction was seen to be associated with *MEF2A* expression (P = 0.03, Figure 3), where the Ndiet-LRFI parental group progeny bulls had

higher expression than all other groups. In SM muscle, RFI parental group was associated with the expression of *MEF2A* (P = 0.01, Figure 2), where bull progeny of the LRFI parents displayed higher expression than the HRFI group. None of the other genes tested in LT and SM muscles were significantly different with respect to RFI or parental diet groups, nor their interaction (Table 6 and 7), although we have seen some potential trends (P < 0.1). In addition, MWT was significantly correlated with mRNA abundance of *IGF2R* (r = -0.42 (transformed *IGF2R*), P = 0.039), whereas FW and TW were positively correlated with mRNA abundance of calpastatin [*CAST*] (FW: r = +0.45, P = 0.026; and TW: r = +0.51, P = 0.009) in LT muscle. In SM muscle, ADG was correlated with mRNA abundance of *CAST* (r = -0.43 (transformed CAST), P = 0.028). As well, FW and TW were significantly correlated with mRNA abundance of myogenin [*MYOG*] (FW: r = +0.41, P = 0.040; and TW: r = +0.42, P = 0.036 (both correlations were performed with transformed *MYOG*)).

Similarly, in testis, RFI parental group was associated with the expression of *MEF2A* (P=0.01, Figure 2), where bull progeny of LRFI parents displayed higher expression than the HRFI parental group. None of the other genes tested in testis were significantly different with respect to parental RFI or maternal diet groups, nor their interaction (Table 8). Regarding the correlation analysis, STDDMI was significantly negatively correlated with expression of insulin like growth factor 1 [*IGF1*] (r = -0.48, P = 0.016), insulin like growth factor 2 [*IGF2*] (r = -0.53, P = 0.007), *IGF2R* (r = -0.57, P = 0.003), and *H19* (r = -0.45, P = 0.024). MWT was significantly negatively correlated with expression of *IGF1* (r = -0.42, P = 0.037), *IGF2R* (r = -0.41, P = 0.043), calpain 1 [*CAPN1*] (r = -0.44, P = 0.028), and *H19* (r = -0.44, P = 0.030). Moreover, ADG, FW, and TW were negatively correlated with mRNA abundance of *IGF2* (ADG: r = -0.39, P = 0.044; FW: r =

-0.50, P = 0.010; and TW: r = -0.42, P = 0.036). As well, FW was negatively correlated with mRNA abundance of *IGF1* (r = -0.40, P = 0.049) and *IGF2R* (r = -0.45, P = 0.023).

4.4. DISCUSSION

In the study of Johnson et al. [73], which was conducted on the same bulls as this study, there was a tendency ($P < 0.1$) for a prenatal maternal diet×time interaction for bull weights, with bulls born from Ldiet fed heifers growing faster between 10 and 16 months of age than bulls born from Ndiet fed heifers. As well, LRFI parental group bulls had a smaller scrotal circumference ($P < 0.01$) and attained puberty ($P < 0.05$) later than HRFI parental group bulls. In the present gene expression study, we wanted to explore the underlying molecular differences of these phenotypes. We also wanted to explore genes previously associated with RFI and prenatal diet [161, 162]. Therefore, we chose to investigate SM and LT muscles, testis and liver to test genes associated with growth and development as well as immunity, to understand some of the underlying molecular biological differences of progeny exposed to different prenatal nutritional treatments in cattle, and genetic selection for divergent RFI.

In this study, only *MEF2A* was differentially expressed in all tissues tested when comparing low and high RFI parental groups. As well, *MEF2A* and *PCDH19* were differentially expressed in LT and liver tissues respectively, when comparing normal versus low maternal diet bulls. We speculate that the low sample number in this study (8 Ndiet-HRFI, 9 Ldiet-HRFI, 3 Ndiet-LRFI, 5 Ldiet-LRFI) affected our ability to identify significant different interactions between prenatal maternal diet and RFI parental groups in gene expression, since we do see these trends in our data. We do feel confident that we have enough individuals to test the main effects, prenatal maternal nutrition as well as parental RFI group, adequately within the confines of this population.

One of the main roles of the liver is to filter the nutrient-rich blood coming from the intestines through the hepatic portal vein, before transferring it to the rest of the body. In particular, the liver detoxifies and breaks down harmful chemicals and this contributes to its active role in immune function [260, 261]. In our study, *PCDH19* was the only immunity associated gene which was differentially expressed, having greater expression in the liver of Ldiet bulls. Zhang et al. [268] reported that *PCDH19* was hyper-methylated and its expression was down-regulated in hepatocellular carcinoma tissues compared to non-tumor tissues in humans. This suggests that expression of *PCDH19* is under epigenetic control, which could make it more susceptible to epigenetic effects of differential prenatal diets, leading to the differences in gene expression detected in this experiment. *PCDH19* expression has also been positively associated with phenotypic measures related to feed intake and efficiency such as daily DMI, MWT, feed conversion ratio (FCR), and RFI in a few studies in the liver of beef cattle [160, 162, 254, 269]. In our study, expression of *PCDH19* was positively correlated with STDDMI and MWT, similar to that reported between the expression of this gene in liver and DMI by Paradis et al. [162], as well as RFI and MWT by Al-Husseini et al. [160]. The lack of differential expression due to the RFI parental group in our study could have been due to the fact that *PCDH19* expression is more strongly associated with phenotypic measures of feed intake rather than genetic potential for RFI (i.e. it is a downstream effect of differences in feed intake rather than an upstream cause). Since there was a trend for higher growth in the Ldiet bulls [73], and our observation that *PCDH19* was positively correlated to STDDMI and MWT, it seems likely that Ldiet bulls ate slightly more than Ndiet bulls to support their increased growth, and in association, had higher *PCDH19* expression in their livers.

Regarding our other correlation results in liver tissue, *GSTM1/2* was positively correlated with RFIf and STDDMI, in agreement with studies of Al-Husseini et al. [160], Chen et al. [254, 269], and de Las Heras-Saldana et al. [270]. The glutathione S-transferase mu 1 [*GSTM1*] and glutathione S-transferase mu 2 [*GSTM2*] family of genes belong to the xenobiotic metabolism signalling pathway and neutralize a variety of toxics, drugs, carcinogens, and products of oxidative stress, through conjugation with glutathione [271]. Oxidative stress occurs when production of reactive oxygen species exceeds the biological system's ability (antioxidant capacity) to detoxify the reactive intermediates or to repair the resulting damage [272]. A recent study conducted by Casal et al. [241] reported lower level of reactive oxidized species in the liver of LRFI steers and thereby lower hepatic oxidative stress compared to HRFI steers. Decreased oxidative stress in liver is associated with lower maintenance requirements, due to a lower lipid and protein turnover and better efficiency in the use of energy [241] Hence the positive relationship between *GSTM1/2* expression with RFIf in liver suggests greater hepatic expression of *GSTM1/2* to catalyze higher rate of reactive oxygen species in cattle with increasing positive values for RFI. Similar to this relationship, increased hepatic expression of *GSTM1/2* may be directly a product of increased STDDMI in higher RFI animals, as the liver would need to be more metabolically active to process higher feed intakes, and in turn process a greater quantity of products of oxidative stress. There was a significant relationship between ADG and MWT with hepatic expression of *H19* and *IGF2R* respectively, although the exact nature of the relationship was ambiguous, since both *H19* and *IGF2R* expression data were transformed before analysis. There was also a negative correlation between ADG and *IGF1R* expression. *H19* is an imprinted gene which has a role as negative regulator (limiting factor) of body weight and cell proliferation [273], via controlling the expression of its adjacent imprinted gene, *IGF2* [274]. The encoded protein of *IGF2* regulates cell

proliferation, growth, migration, differentiation and survival via binding to its receptors such as IGF1R (IGF2 increases cell proliferation, growth, etc. via binding to IGF1R) and IGF2R (IGF2 is rendered ineffective when it binds to circulating IGF2R) [275]. Unlike *IGF1,* which is a major growth factor in adults, *IGF2* is highly active during fetal development and less active after birth, and its exact role in the adult is not fully understood. In adults, *IGF2* expression occurs mostly in liver and in epithelial cells lining the surface of the brain [275]. Our observation that expression of *H19*, *IGF2R*, and *IGF1R* in liver was associated with growth traits in these bulls, albeit in an unexpected manner (i.e. the negative correlation of *IGF1R* with the growth related trait ADG), gives some insight on the possible interplay between these genes and their physiological effects in the postnatal animal.

MEF2A is an important gene involved in muscle development [276, 277]. It is a positive regulator in skeletal muscle myoblast proliferation possibly via regulating myozenin 2 [*MyoZ2*] [277]. The protein encoded by *MEF2A* is a member of the myocyte enhancer factor 2 (MEF2) protein family which is involved in vertebrate skeletal, cardiac, and smooth muscle development, as well as differentiation during myogenesis [276]. Thus, *MEF2* genes are considered as major regulators of postnatal skeletal muscle growth [276]. In the comparison between HRFI and LRFI parental group bulls, *MEF2A* was significantly differentially expressed in liver, SM muscle, and testis, with consistently greater expression in the LRFI parental group bulls (Figure 2). *MEF2A* was also higher in the average expression of LRFI parental group bulls in LT muscle (Figure 3). *MEF2A* expression in all the tissues tested was not significantly correlated with phenotypic RFIf measured on progeny bulls while they were in the GrowSafe test (data not shown). Therefore, we speculate that the presence of polymorphism(s) in the promoter/coding region of *MEF2A* is linked to genotypes associated with RFI in our Angus population, but may not be causative of the RFIf

phenotype. Juszczuk-Kubiak et al. [276] reported effects of single nucleotide polymorphisms (SNPs) at position -780T/G, -768T/G, and -222A/G in the promoter region of *MEF2A*. They detected higher expression of *MEF2A* in the *Longissimus dorsi* (LD) muscle of bulls carrying the homozygous TTA/TTA genotype rather than those with heterozygous TTA/GGG variant. Therefore, polymorphisms exist that affect the level of *MEF2A* expression, which could be linked to our low and high RFI parental groupings. To that end, others in our research group have detected three SNPs upstream of *MEF2A* to be significantly associated with RFI in a 7.8 million SNP genome-wide association study (GWAS) performed in our Angus population, but other SNPs in the population had much larger effects upon RFI and associated traits (Changxi Li, personal communication).

Expression of *MEF2A* in LT muscle was associated with a maternal prenatal diet×RFI parental group interaction (Figure 3), where its expression was significantly higher in Ndiet-LRFI parental group bulls compared to other three groups. This result represents an example of a genotype×environmental effect upon *MEF2A* expression in LT muscle. As discussed earlier, expression of *MEF2A* appears to be linked to parental RFI genetic potential, with LRFI parental group bulls displaying higher expression. Higher expression of *MEF2A* in LT muscle of Ndiet-LRFI parental group bulls over all groups might be plausible, since the nutritionally unrestricted dams would have consumed enough nutrients to partition energy towards LT muscle in the fetus and further increase *MEF2A* expression over the Ldiet-LRFI parental group bulls. The other tissues tested may not have shown this pattern so dramatically, because they are more essential to fetal survival - i.e., the liver, which is an essential organ, and SM, which is important for postnatal locomotion [3-5, 96]. A similar pattern of differential effects upon various fetal tissues from prenatal nutrition in bovine dams was reported by Paradis et al. [161]. In that study, there was a

higher level of gene expression differences in *Longissimus dorsi* muscle of high versus low maternal diet fetuses than in semitendinosus, a muscle with similar functions to SM. Testis tissue may not have reflected the maternal diet effect in LT on the Ndiet-LRFI group more than the others, since gonadal mitotic activity is lower during the period of 65-230 days of gestation [278], which was when our maternal dietary treatment took place, or that differential regulation of *MEF2A* expression by maternal diet in testis would gain no advantage for testicular development since the gene is associated with muscle development.

The mRNA abundance of *IGF2R* in LT muscle was correlated with MWT, plus the mRNA abundance of *MYOG* in SM muscle was correlated with FW and TW, although the exact nature of the relationship is ambiguous, since both *IGF2R* and *MYOG* expression data were transformed before analysis. The primary role of *IGF2R* is in hydrolysis of circulating *IGF2* [279], which is an important fetal mitogen [280]. *MYOG* is a member of myogenic regulatory factors (MRF) and is involved in embryonic skeletal muscle fiber differentiation (permits the transition from proliferating myoblasts to differentiating myotubes) [281]. In adults, *MYOG* controls the balance between muscle hypertrophy and muscle atrophy which are essential for maintenance of skeletal muscle mass [282]. Therefore, any impacts upon muscle development in utero may also affect the expression of both *IGF2R* and *MYOG* postnatally. As a result, it is plausible that their expression may continue to be related to growth traits in these bulls, particularly their expression in muscle tissue.

FW and TW were positively correlated with mRNA abundance of *CAST* in LT muscle. The calpastatin protein encoded by *CAST*, is a specific inhibitor of the calpain family and is involved in *in vivo* cell proliferation and cell viability [283]. The calpain family proteins including μ-calpain and m-calpain are implicated in myofibrillar protein degradation [284]. Hence, the positive

correlation between body weight traits and expression of *CAST* was plausible, as its increased expression reduces a positive regular of muscle protein degradation. In SM muscle, ADG was also correlated with mRNA abundance of *CAST*, although the exact nature of the relationship was ambiguous, as *CAST* expression data were transformed before analysis. Morgan et al. [285] reported a significant negative correlation between calpastatin activity and fractional degradation rates (FDR) of *Longissimus* muscle. We inferred that decreased FDR of skeletal muscle, possibly through higher calpastatin activity, was associated with greater efficiency of growth [285]. Studies of Nassiri et al. [286], Chung and Davis [287], as well as Khan et al. [288] have also reported an association between ovine *CAST* polymorphisms and ADG. As well, Karisa et al. [145] found two SNPs at 271G/A and 672A/G positions of the *CAST* gene, which were associated with RFI, in beef steers. Therefore, it is plausible that genetic variation in our population might exist that affected the level of *CAST* expression, which could be linked to body weight (FW and TW) and phenotypic measures of RFI (MWT and ADG) in our study.

With regards to the other significant negative correlations between gene expression in testis and phenotypic measures of intake (STDDMI with *IGF1*, *IGF2*, *IGF2R*, and *H19*), and growth (ADG, FW, and TW with *IGF2*; FW with *IGF1* and *IGF2R*; as well as MWT with *IGF1*, *IGF2R*, *CAPN1*, and *H19*), the detailed biological origins of these associations were uncertain. The testes have two main functions including: 1) production of haploid germ cells (spermatogenesis) necessary for reproduction, and 2) synthesis of androgens (primarily testosterone) which are necessary for spermatogenesis and play an important role in the development of male sex characteristics [289]. However, as recently reported by Uhlen et al. [262], testis tissue expresses the highest number of genes compared to other tissues in human body. Therefore, expression of some key genes involved in growth and development including *IGF1*, *IGF2*, *IGF2R*, *CAPN1*, and

H19 in testis and their association with phenotypic measures of intake and growth could be a reflection of how these genes are expressed in other tissues, ultimately resulting in modulation of growth and development, with less of an influence of daily biological variation between individuals due to environmental impacts.

Other studies have found differential expression due to both prenatal maternal diet and RFI group for some of the other genes that we have tested [161, 162, 171, 193, 290], although several differences existed in the breed and sex of the animals selected for the high and low RFI groups, the type, time window, and severity of dietary treatment, age and breed of cattle selected for the dietary treatment, and age of animal during tissue collection, which could all lead to differences in the results of our studies. This is true for nearly all prenatal nutritional studies in cattle, which makes it extremely difficult to confirm or refute the results of any single study. It was also reported by Mukiibi et al. [291] that gene expression profiles in the liver of steers divergent for RFI can be highly breed-specific. Bearing that in mind, the accumulation of studies and information about the effects of prenatal nutrition on postnatal growth and development in cattle, and genes that are differentially expressed between high vs. low RFI cattle, can start to bear fruit in the identification of biological pathways and processes that are most likely to be affected, and give insights into methods to optimize prenatal nutrition to ensure the health and prosperity of the postnatal animal.

4.5. CONCLUSIONS

In conclusion, expression of *MEF2A* was related to parental group RFI in all tissues tested, as well as prenatal maternal nutrition in LT. Furthermore, *PCDH19* expression in liver tissue was most likely related to feed intake differences in normal versus low maternal diet bulls. Therefore, the phenotypic differences seen in these bulls were most likely manifested at an earlier age, and/or are not related to many of the genes we chose to study. Although many of the genes surveyed in this

study did not show differential expression due to parental RFI or maternal diet, one gene (*GSTM1/*2) showed strong associations with RFIf in liver, plus some (*IGF1*, *IGF1R*, *IGF2*, *IGF2R*, *H19*, *PCDH19*, *CAPN1*, *CAST*) showed significant associations with phenotypic measures of RFI (STDDMI, ADG, and/or MWT) in different tissues. These findings help to understand the underlying biological mechanisms regulating postnatal responses to prenatal nutrition, and feed efficiency, intake, and growth traits in beef bulls.

Table 4. 1 Nutrient analysis of ration on a dry matter basis (%DM) fed to heifers during the initial GrowSafe trial, 30-150 days of treatment and 150 days to parturition

Diet composition (%DM)	**GrowSafe trial ration**	**30-150 days ration**		**150 days to parturition**
Feed type	TMR[1]	Brome grass hay	Oats	Hay
DM[2] % (actual)	62.0	81.8	89.7	81.4
CP[3] (%DM)	11.9	10.3	11.9	18.7
ADF[4] (%DM)	28.6	49.5	15.0	32.3
NDF[5] (%DM)	45.7	70.1	29.9	39.1
TDN[6] (%DM)	67.5	52.8	77.0	64.8
Ca (%DM)	1.67	0.65	0.13	1.67
P (%DM)	0.39	0.25	0.37	0.26
Mg (%DM)	0.23	0.20	0.15	0.33
K (%DM)	0.91	2.55	0.59	2.12
Na (%DM)	0.42	0.01	0.02	0.04
Salt (%DM)	1.05	0.04	0.04	0.10

[1] TMR: Total Mixed Ration; [2] DM: Dry Matter Basis; [3]CP: Crude Protein; [4] ADF: Acid Detergent Fibre; [5] NDF: Neutral Detergent Fibre; [6] TDN: Total Digestible Nutrients

Table 4. 2 Dietary adjustment of normal and low diets over the course of the trial (as fed)

Adjustments	**Normal**		**Low**	
	Hay (kg/day)	**Oats (kg/day)**	**Hay (kg/day)**	**Oats (kg/day)**
Initial [1]	6.94	3.45	9.71	0.00
1st adjustment [2]	7.73	3.75	9.33	0.00
2nd adjustment [3]	9.21	3.95	11.29	0.00
3rd adjustment[4]	7.14	5.65	12.84	0.00
4th adjustment[5]	7.02	6.80	10.78	2.61

[1] Initial: Jul. 24, 2013; [2] 1st adjustment: Aug. 19, 2013; [3] 2nd adjustment: Sept. 11, 2013; [4] 3rd adjustment Oct. 8, 2013; [5] 4th adjustment: Nov. 8, 2013

Table 4. 3 Nutrient analysis of barley-silage ration fed to bulls during the RFI test period in the GrowSafe system

Diet composition	Value
DM[1] % (actual)	56.1
CP[2] (%DM)	14
ADF[3] (%DM)	25.3
NDF[4] (%DM)	40.5
TDN[5] (%DM)	69.6
Ca (%DM)	0.94
P (%DM)	0.34
Mg (%DM)	0.23
K (%DM)	1.38
Na (%DM)	0.13
Fe (PPM)	336
Mn (PPM)	70
Zn (PPM)	61
Cu (PPM)	16

[1] DM: Dry Matter Basis; [2] CP: Crude Protein; [3] ADF: Acid Detergent Fibre; [4] NDF: Neutral Detergent Fibre; [5] TDN: Total Digestible Nutrients

Table 4. 4 Forward and reverse primers for endogenous control genes

Gene name	Gene symbol	Forward Primer (5'-3')	Reverse Primer (5'-3')	GenBank Accession number	Tissue	Product Size (bp)	Association
Ribosomal protein L19	*RPL19*[E]	ACCCCAATGAGACCAATGAA	ATGGACAGTCACAGGCTTCC	NM_001040516.1	All[1]	101	Housekeeping gene
Beta actin	*ACTB*[E]	CTCTTCCAGCCTTCCTTCCT	CCAATCCACACGGAGTACTTG	NM_173979.3	All	245	Housekeeping gene
Peptidylprolyl isomerase A (cyclophilin A)	*PPIA*[E]	GTCAACCCCACCGTGTTCT	TCCTTTCTCTCCAGTGCTCAG	NM_178320.2	All	132	Housekeeping gene
Eukaryotic translation elongation factor 1 alpha 2	*EEF1A2*[E]	AGTTCACGTCCCAGGTCATC	CTCCAACTTCTTGCCAGAGC	NM_001037464.1	All	149	Housekeeping gene
Glyceraldehyde-3-phosphate dehydrogenase	*GAPDH*[E]	TGACCCCTTCATTGACCTTC	GATCTCGCTCCTGGAAGATG	NM_001034034.1	All	143	Housekeeping gene
Hydroxymethylbilane synthase	*HMBS*[E]	CTACTTCGCTGCATTGCTGA	CAGGTACAGTTGCCCATCCT	NM_001046207.1	All	105	Housekeeping gene
Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta	*YWHAZ*[E]	AGACGGAAGGTGCTGAGAAA	CGTTGGGGATCAAGAACTTT	NM_174814.2	All	123	Housekeeping gene
Maternally expressed 3 (non-protein coding)	*MEG3*[I]	TCACCTGTCTCACGCTTCTC	GACCAGAGGAGACCACGAAG	NR_037684.1	All	171	Immunity
Insulin like growth factor 2	*IGF2*[I]	CCAGCGATTAGAAGTGAGCC	AGACCTAGTGGGGCGGTC	NM_174087.3	All	95	Muscle development
Insulin like growth factor 2 receptor	*IGF2R*[I]	GCAATGCTAAGCTTTCGTATTACG	GGTGTACCACCGGAAGTTGTATG	NM_174352.2	All	188	Muscle development
H19 imprinted maternally expressed transcript	*H19*[I]	CAGACACACCACTGCTGCTC	GAAGTCCGTGTTCCAAGTCC	NR_003958.2	All	97	Immunity, growth and development
Neuronatin	*NNAT*[I]	AGCACTCGCTCTCAACCAC	GGAAAATGTACCAGCCGATG	NM_001201324.1	All	98	Immunity
Insulin like growth factor 1	*IGF1*[N]	GATGCTCTCCAGTTCGTGTG	CTCCAGCCTCCTCAGATCAC	NM_001077828.1	All	141	Immunity and muscle development
Insulin like growth factor 1 receptor	*IGF1R*[N]	CAAAGGCAATCTGCTCATCA	CAGGAAGGACAAGGAGACCA	NM_001244612.1	All	139	Immunity and muscle development
Myogenic differentiation 1	*MYOD1*[N]	GAACACTACAGCGGCGACTC	AGTAAGTGCGGTCGTAGCAG	NM_001040478.2	All	121	Muscle development
Myogenin	*MYOG*[N]	CAGTGAATGCAGCTCCCATA	CGACATCCTCCACTGTGATG	NM_001111325.1	All	164	Muscle development
Myocyte enhancer factor 2A	*MEF2A*[N]	CAATGCCAACTGCCTACAAC	TGTCCTAAATGGTGCTGCTG	NM_001083638.2	All	130	Muscle development
Calpain 1	*CAPN1*[N]	CCTGCTGGAGAAAGCCTATG	GCTCGTACCACTCGGTGACT	NM_174259.2	All	113	Muscle development
Calpastatin	*CAST*[N]	GTGCCTCGGACCTCTATGTG	CGTCTTCTGGATCTGCTTCC	NM_001030318.3	All	133	Muscle development
ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit	*ATP5O*[N]	GAAGGAGTTGTTGCGAGTAGG	TTGCCGTCATGTCACTTAGG	NM_174244.1	Liver	116	Immunity
Glutathione S-transferase M1 and M2	*GSTM1/2*[N]	GGGAGACAGAGGAGGAGATGA	CCTTCAAGAAACCAGGCTTCA	NM_175825.3	Liver	126	Immunity
Solute carrier family 2 member 1	*SLC2A1*[N]	ACACAGCCTTCACTGTCGTG	TGCTCAGGTAGGACATCCAG	NM_174602.2	Liver	156	Immunity
Interferon-induced protein 44 (IFI44), transcript variant X1	*IFI44*[N]	ACGCATGTGGATACCTTGGA	AGGACATCTATGACAGGCTCC	XM_002686295.3	Liver	179	Immunity
Insulin like growth factor binding protein 3	*IGFBP3*[N]	CCTCTGAGTCCAAGCGTGAG	GCTGCCCGTACTTATCCACA	NM_174556.1	Liver	210	Immunity and muscle development
Inhibin beta A subunit	*INHBA*[N]	GGACGGAGGGCAGAAATGAA	AGACGGATGGTGACTTTGCT	NM_174363.2	Liver	203	Immunity
Protocadherin 19	*PCDH19*[N]	GAACACCAGTGTGACCTCCA	GCTTCAACATCAGCAGCAGT	XM_003588123.2	Liver	207	Immunity
MX dynamin like GTPase 1	*MX1*[N]	TTCAACCTCCACCGAACTGC	TGCCTCCTTCTCTCTGACCT	NM_173940.2	Liver	165	Immunity
Metallothionein 1E	*MT1E*[N]	CAACTGCTCCTGCTCCACT	CCCACGTTCCTCCATTGATA	NM_001114857.2	Liver	221	Immunity
ISG15 ubiquitin-like modifier	*ISG15*[N]	CGCAGCCAACCAGTGTCT	CGTCATGGAGTCCCTCAGA	NM_174366.1	Liver	120	Immunity
Heat shock 60kDa protein 1	*HSPD1*[N]	TCCAATCCATTGTTCCTGCT	CTGCCACAACTTGAAGACCA	NM_001166608.1	Liver	138	Immunity
Hect domain and RLD 6	*HERC6*[N]	GTTCCACCAGTGTTCCCAGG	GCAGTCAGACAAGCAGGAGA	NM_001192644.1	Liver	157	Immunity
Heat shock protein family A (Hsp70) member 5	*HSPA5*[N]	TGAAACTGTGGGAGGTGTCA	CCAGAAGGTGATTGTCTTTCG	NM_001075148.1	Liver	161	Immunity
Toll like receptor 4	*TLR4*[N]	CAGAACAACTTGCTCCCTGAC	GCACCTGAAGGCTAGAGAGG	NM_174198.6	Liver	124	Immunity

[1] LT muscle, SM muscle, liver and testis tissues; [E] Endogenous gene; [I] Imprinted gene; [N] Non-imprinted gene

Table 4. 5 Effect of prenatal diet and RFI on gene expression of liver of progeny bulls

	P-value		
Gene Symbol	**Diet**	**RFI**	**Diet×RFI**
IGF1	0.64	0.26	0.42
IGF1R	0.44	0.97	0.61
IGF2	0.97	0.97	0.18
IGF2R	0.77	0.7	0.42
MEF2A	0.19	0.01*	0.09
CAPN1	0.23	0.15	0.91
CAST	0.19	0.15	0.53
H19	0.29	0.45	0.49
HSPD1	0.59	0.48	0.63
PCDH19	0.01*	0.3	0.16
GSTM1/2	0.12	0.09	0.86
ATP5O	0.91	0.51	0.28
MX1	0.99	0.88	0.89
IGFBP3	0.74	0.42	0.6
HSPA5	0.08	0.16	0.55
TLR4	0.26	0.39	0.87

* $P < 0.05$

Table 4. 6 Effect of prenatal diet and RFI on gene expression of LT muscle of progeny bulls

	P-value		
Gene Symbol	**Diet**	**RFI**	**Diet×RFI**
IGF1R	0.21	0.27	0.56
IGF2	0.35	0.74	0.34
IGF2R	0.07	0.81	0.52
MEF2A	0.02*	0.01*	0.03*
CAPN1	0.67	0.16	0.92
CAST	0.17	0.29	0.96
H19	0.45	0.11	0.51
MYOD1	0.35	0.39	0.87
MYOG	0.52	0.28	0.76

* $P < 0.05$

Table 4. 7 Effect of prenatal diet and RFI on gene expression of SM muscle of progeny bulls

	P-value		
Gene Symbol	**Diet**	**RFI**	**Diet×RFI**
IGF1	0.92	0.17	0.66
IGF1R	0.09	0.1	0.83
IGF2R	0.87	0.97	0.91
MEF2A	0.33	0.01*	0.16
CAPN1	0.78	0.13	0.66
CAST	0.22	0.14	0.44
H19	0.32	0.13	0.8
MYOD1	0.48	0.98	0.19
MYOG	0.82	0.45	0.37

* $P < 0.05$

Table 4. 8 Effect of prenatal diet and RFI on gene expression of testis of progeny bulls

	P-value		
Gene Symbol	**Diet**	**RFI**	**Diet×RFI**
IGF1	0.94	0.59	0.09
IGF1R	0.44	0.48	0.48
IGF2	0.69	0.26	0.3
IGF2R	0.15	0.76	0.64
MEF2A	0.1	0.01*	0.52
CAPN1	0.8	0.16	0.33
CAST	0.89	0.57	0.7
H19	0.17	0.92	0.73
NNAT	0.78	0.85	0.5

* $P < 0.05$

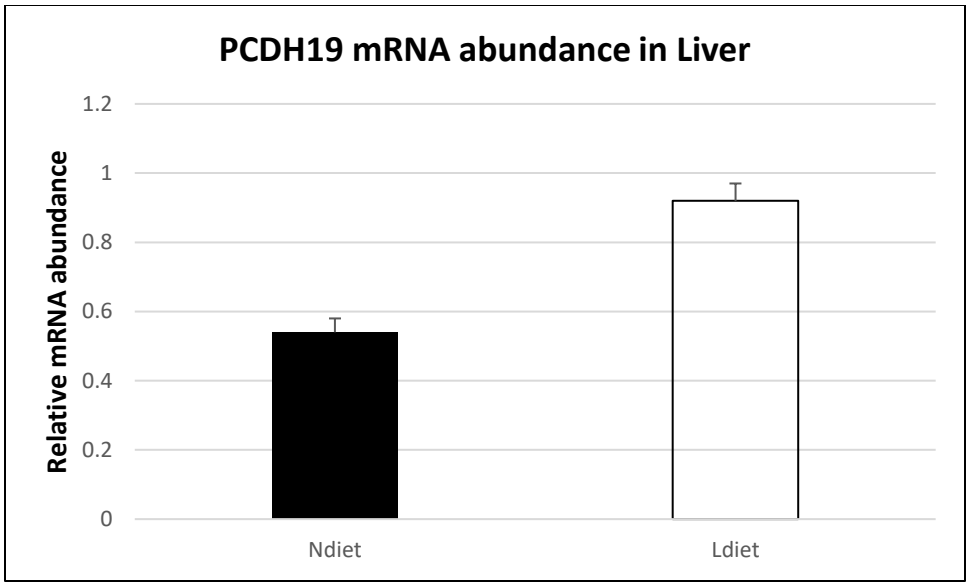

Figure 4. 1 Comparison of mRNA abundance of *PCDH19* gene in liver of Ndiet vs. Ldiet groups. Data represent significant ($P < 0.05$) differences between Ndiet and Ldiet groups within *PCDH19* gene and are expressed as least squares means ± SEM.

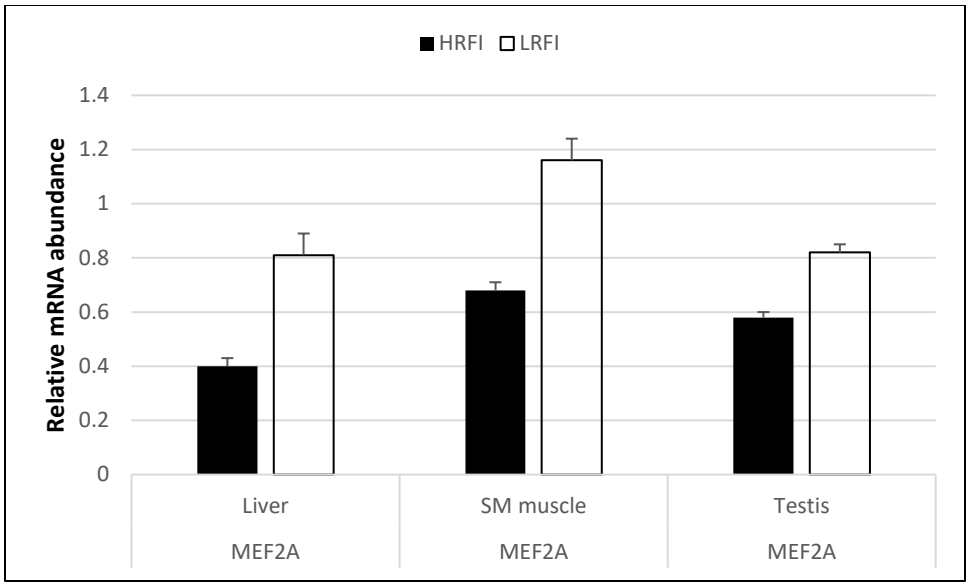

Figure 4. 2 Comparison of mRNA abundance of *MEF2A* gene in liver, SM muscle, and testis of HRFI vs. LRFI groups. Data represent significant ($P < 0.05$) differences between RFI groups within *MEF2A* gene and are expressed as least squares means ± SEM.

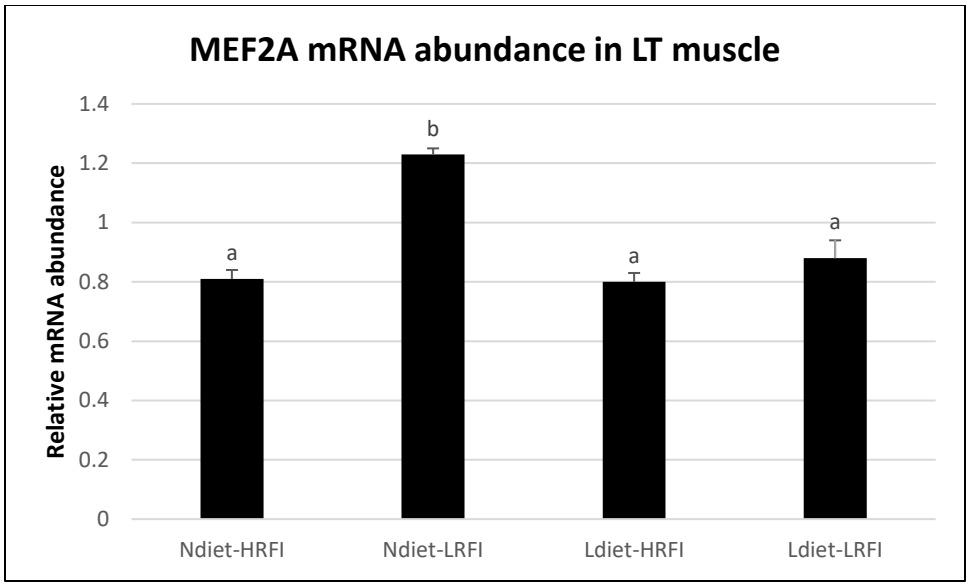

Figure 4. 3 Comparison of mRNA abundance of *MEF2A* gene in LT muscle by dietary treatment and RFI. Data are expressed as least squares means ± SEM. [a,b] Different letters represent significant differences between treatment groups ($P < 0.05$).

CHAPTER 5. IMPACT OF PRENATAL MATERNAL NUTRITION AND SELECTION FOR PARENTAL RESIDUAL FEED INTAKE (RFI) ON SPERM DNA METHYLATION PATTERNS IN YOUNG ANGUS BULLS

5.1. INTRODUCTION

DNA methylation/demethylation provides a crucial layer of epigenetic regulation for embryogenesis and normal development by permitting the modification or reprogramming of gene expression [292-296]. The incomplete erasure of DNA methylation marks during the reprogramming process allows transfer of some environmental impacts from one generation to the next [297]. Several studies have reported that aberrant sperm DNA methylation, due to alterations in environmental exposures (i.e. air pollution, lead, mercury, tobacco smoke, and nutritional factors), can cause impaired male fertility, alter embryonic development and increase offspring's susceptibility to various disorders, i.e. neuropsychiatric and fecundity disorders [298-303]. This highlights the importance of sperm DNA methylation in the fecundity, health, and development of offspring.

Alterations in prenatal nutrition experienced by parents can cause new epigenomic marks to appear in their germline genome, which can be passed to the next generations and thereby affect the phenotypes of their offspring, especially susceptibility to disease in later life [304]. Heijmans et al. [301] reported that individuals who were prenatally exposed to famine during the Dutch Hunger Winter had less DNA methylation of the imprinted *IGF2* gene compared with their unexposed, same-sex siblings. This loss of methylation was associated with reduced birth weight, as well as a higher tendency to develop obesity and later-life health issues. In livestock, a number of studies have shown the effects of maternal dietary perturbations on offspring traits such as

postnatal growth and development, fecundity, and general health [3, 4, 85, 305]. For example, several studies have shown that nutrient insults during the last two trimesters of pregnancy are associated with reduced fetal growth and birth weight of the offspring in both sheep and cattle [87, 89, 90]. At the molecular level, a study by Lan et al. [192] reported that the methylation levels of CpG islands of both the *H19* [H19 fetal liver mRNA] gene and the *IGF2R* [insulin like growth factor 2 receptor] gene were higher in the fetuses of pregnant sheep that were fed alfalfa haylage (fiber) and dried corn distiller's grains (fiber plus protein plus fat) than those fed only corn (starch). Paradis et al. [161] also reported that feed restriction of 85% compared with 140% of total metabolizable energy requirements during mid-to-late gestation could alter the expression of growth, adipogenic and myogenic genes in fetal bovine muscle without apparent differences in phenotype. These data provide substantial evidence of detectable epigenomic marks in offspring, arising from maternal dietary differences, which results in differences in gene expression and thereby metabolism alterations in different parts of body, leading to epigenetic programming.

A genetic trait that is becoming popular for selection in beef cattle, which could potentially modulate prenatal nutrition effects upon the unborn offspring, is residual feed intake (RFI). RFI is a feed efficiency measure defined as the difference between an animal's actual feed intake and its expected feed requirements for maintenance and growth over a specific time period [14, 16]. RFI is an appealing measure of feed efficiency because it is moderately heritable (h^2 = 0.29-0.46) in cattle and is independent of growth characteristics such as body weight (BW) and average daily gain (ADG) [16, 22, 28]. In beef cattle, selection for low-RFI (LRFI) animals (feed efficient animals) is gaining popularity among beef producers as these animals eat less and produce less methane per unit weight gain [7, 222]. Similar to prenatal maternal diet, the RFI status of an animal is also associated with metabolism alterations in different parts of body [30, 306].

In terms of bull fertility, several studies have shown that scrotal circumference (SC) and semen characteristics were not statistically different between LRFI and high-RFI (HRFI) bulls [76, 307-309]. In contrast, other studies found that LRFI bulls had smaller SC [73], decreased sperm motility [73, 76, 77], decreased progressive sperm motility, increased abundance of tail abnormalities, and delayed sexual maturity [78, 79]. The mechanisms behind the negative associations are unclear. Both negative genomic associations between RFI and fertility parameters, and possibly epigenomic marks may also play a key role during this process. In cattle, only one study to date has focused on associations between RFI and DNA methylation [191]. They identified 1493 differentially methylated cytosines (DMCs) and 279 differentially methylated regions (DMRs) in the hepatic tissue of Nelore cattle that were characterized as extremes for RFI. Hence, methylation patterns of DNA may be associated with RFI and therefore it is important to study this relationship.

The biological mechanisms by which prenatal maternal diet and/or RFI influence the offspring's phenotype, or subsequent generations' phenotype, remains obscure. As a result, this study was undertaken to identify DMRs (via whole genome bisulfite sequencing - WGBS) in the sperm of progeny Angus bulls whose dams were exposed to different prenatal maternal diets during first half of gestation. In addition, we also aimed to identify DMRs in the sperm of progeny Angus bulls with divergent parental RFI that were exposed to different prenatal diets (low-diet (Ldiet) and normal-diet (Ndiet)). The identified DMRs were further interrogated to identify the biological pathways associated with RFI and prenatal diet, which were manifested through sperm epigenomics.

5.2. MATERIALS AND METHODS

5.2.1. Ethics approvals

The collection and analysis of bovine semen in this study were approved by the University of Alberta's Animal Care Committee (Animal Use Protocol [AUP] 1129) under the auspices of the Canadian Council of Animal Care [209].

5.2.2. Animals and experimental design

Details of heifer selection, experimental design, estrous-synchronization and artificial insemination (AI), as well as prenatal dietary treatment were reported in detail by Johnson et al. [73], and in chapter 4 of this thesis. Briefly, 89 purebred Angus heifers raised on the University of Alberta's Roy Berg Kinsella Research Ranch (Kinsella, Alberta, Canada), were used for this study. The population in which these heifers originated from was described previously by Mao et al. [210]. During late winter and early spring of 2013, heifers were tested for RFI at approximately 9 to 12 months of age using the GrowSafe™ automated feed recording system (GrowSafe Systems Ltd., Airdrie, Alberta, Canada) as described by Basarab et al. [16] and Mao et al. [210]. After the test, they were classified as either HRFI or LRFI by ordering the heifers based upon their corresponding measurement of RFI corrected for backfat [37], from negative to positive, and then dividing them in half. To produce calves with divergent potential for RFI, heifers classified as HRFI were bred to sires identified as HRFI (n=2), and heifers classified as LRFI were bred to sires identified as LRFI (n=2). Consequently, the progeny of these matings were expected to have divergent genetic potentials for RFI because of the assortative breeding scheme. At the time of selection, the RFI estimated breeding values (EBVs) of the bulls were +0.174 and +0.140 kgs DMI/day (HRFI), and -0.230 and -0.482 kgs DMI/day (LRFI), with accuracies of 0.859, 0.874,

0.805, and 0.712 respectively. Bulls were also selected so that EBVs for other recorded traits were similar across all four bulls. Heifers within each RFI group were randomly stratified into two sub-groups, and each subgroup was assigned to one of the bulls within their respective RFI classification (High versus Low) for breeding via estrous synchronization and AI. Between the four heifer sub-groups there were no biases in physical measurements collected during the RFI test, such as end-of-test weight, ADG during test, metabolic mid-weight, and ultrasound backfat (data not shown). Also, there was no difference in standardized dry matter intake nor RFI for heifer groups assigned to AI bull within RFI classification (data not shown).

Breeding was performed via two estrous-synchronized AIs as reported by Johnson et al. [73]. The first AI was timed, and the second was performed 12 hours after the detection of estrous. Pregnancy determination was performed at 28 days after each AI via trans-rectal ultrasonography (Aloka-500V scanner equipped with a 7.5 MHz linear transducer; Aloka Co., Tokyo, Japan), and heifers that did not become pregnant after the two rounds of AI were removed from the study. The remaining pregnant heifers (n = 61) were stratified by RFI, body weight, backfat depth, AI sire, and conception date, and randomly but equally assigned to one of two dietary treatments from day 30 to day 150 of gestation: 1) A low-diet or Ldiet formulated for 0.5 kg/d ADG consisting of Brome Grass hay (~75% of NRC requirements for growing pregnant heifers, and of the normal diet); or 2) a normal-diet or Ndiet formulated for 0.7 kg/d ADG consisting of Brome Grass hay supplemented with oats (roughly 100% of NRC requirements for growing, pregnant heifers). To account for heifer growth and increasing weight of the conceptus, the ration was adjusted approximately once per month after the heifers were weighed. Heifers were fed the entire ration once a day, and if oats were fed it was provided in separate bunks before the hay was fed. Individual feed intake of Brome Grass hay was recorded using the GrowSafe System. Supplemental salt and

minerals were provided free-choice [73]. Two heifers were removed from the prenatal nutrition trial after they were found not to be pregnant due to a subsequent pregnancy check.

After 150 days of gestation, all heifers were housed together and offered free-choice hay until approximately 2 months after birth, when they grazed mixed tame and native grass pasture until weaning in November 2014. The male calves born to these heifers were left as bulls (11 Ndiet vs. 15 Ldiet and 18 HRFI vs. 8 LRFI). After weaning, bulls were fed and managed according to industry standards for production of potential replacement yearling bulls in Alberta (Johnson et al., 2019).

5.2.3. Sample collection

Details of semen collection is reported in detail by Johnson et al. [73]. Briefly, semen was collected by electroejaculation at 4-week intervals from 13 to 16 months. Samples with ≥ 50% motility were diluted with equal volumes of semen extender (AndroMed-Minitube, Ingersoll, ON, Canada) and transferred to 15 mL tubes. Subsequently, the concentration of sperm was determined using a hemocytometer (Electron Microscopy Sciences, Hatfield, PA, USA), which was followed by adjusting the final concentration to 50×10^6 sperm/mL. The semen was then transferred into 0.5 mL straws, frozen using the automatic freezing machine (ICE Cube: Minitube, Ingersoll, Ontario, Canada), and plunged into liquid nitrogen for long-term storage [310].

5.2.4. DNA isolation

The final collection of semen (16^{th} months of age) was selected for analysis. To begin the DNA extraction procedure, individual semen straws were thawed at room temperature and 125 µL of each semen straw was transferred to a 1.5 mL Eppendorf tube. One mL of 1X STE buffer (100mM Tris; pH=8, 10mM EDTA; pH=8, 1M NaCl) was added to the tube, and after a brief vortex, the

tube was centrifuged at 7000 ×*g* for 5 min. The supernatant was poured off and the remaining pellet was washed (to separate clumps of cells from semen extender) again twice with 1 mL 1X STE buffer. After the final wash, the supernatant was again poured off and the pellet was resuspended in 336 µL 1X STE. The tube was vortexed, then 40 µL of 10% SDS (final concentration ~1%), 10 µL of 20 mg/mL Proteinase K, and 14 µL of 1 M DTT (final concentration ~35 mM) were added to the tube. The tube was incubated in a rotating hybridization oven at 56 °C overnight for 24 hrs. Two MaXtract tubes (Qiagen, Hilden, Germany) were prepared by pulse centrifugation at 15,000 ×*g* for 20-30 sec. Then 400 µL of phenol:chloroform:isoamyl alcohol (PCI=25:24:1) was added to one set of MaXtract tubes. The sperm cell lysate from the overnight incubation was poured into the same MaXtract tube, mixed by inversion for 5 min, and centrifuged at 15,000 ×*g* for 5 min at 4ºC. A total of 400 µL of chloroform:isoamyl alcohol (CI=24:1) was added to the other set of MaXtract tubes. The aqueous layer (top) was transferred from the PCI tube to the CI tube, mixed by inversion for 5 min, then centrifuged at 15,000 ×*g* for 5 min at 4 ºC. Then 0.1 volume of 3 M sodium acetate (~40 µL) was added to a new 1.5 mL tube, and 400 µL of the aqueous layer was transferred into the tube containing sodium acetate. Two volumes of 100% ethanol (~880 µL) was added and the tube was frozen at -80 ºC for 2 hrs. Following the freezing process, the tube was left at room temperature for 3 min until the liquid inside melted. After melting, the tube was centrifuged at 15,000 ×*g* for 15 min at 4 ºC. The ethanol supernatant was poured off into a waste flask, and 1 mL of 70% ethanol was added to the pellet, which was followed by centrifugation at 15,000 ×*g* for 5 min at 4 ºC. The pellet was washed with 70% ethanol for a second time and then the pellet was resuspended in 100 µL AE buffer (10 mM Tris-Cl, 0.5 mM EDTA; pH=9.0) (Qiagen). Total DNA was quantified using a Nanodrop-2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, NC, USA) and DNA integrity was evaluated using an

Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA). The sample was then stored at -80 °C. All DNA samples used in this study had a DNA integrity number (DIN) score of ≥ 8.5. Afterwards, DNA samples were shipped to the Beijing Genomics Institute (BGI, Shenzhen, China), in a box containing dry ice, for library construction and whole-genome bisulfite sequencing. BGI remeasured the DNA quantity using a Qubit 3 Fluorometer (Life Technologies, Carlsbad, CA, USA) and 1% Agarose Gel Electrophoresis (voltage: 150 V, electrophoresis time: 40 min) before library construction.

5.2.5. Library construction and whole-genome bisulfite sequencing

One µg of genomic DNA was sheared by sonication (Covaris, Woburn, MA, USA). Shearing was followed by 1% Agarose Gel Electrophoresis, and bands corresponding to DNA insert sizes of 200-300 bp were excised then purified with QIAquick Gel Extraction kit (Qiagen, Germantown, MD, USA). The fragmented DNA was end-repaired, then purified with MiniElute PCR Purification Kit (Qiagen). A single adenosine nucleotide was added at the 3′ extremities of the blunt-end fragments, then purified with MiniElute PCR Purification Kit (Qiagen). Methylated adapters (BGI in-house adapter with length of 124 bp) were ligated to the adenylated 3' ends of each strand in the genomic fragments, then purified with MiniElute PCR Purification Kit (Qiagen). Then DNA fragments were treated with sodium bisulfite using the EZ DNA Methylation-Gold™ Kit (Zymo Research, Irvine, CA, USA). Two rounds of conversion were performed to achieve >99% conversion. Agarose gel electrophoresis was performed on the ligation product, and fragments ranging from 320 to 420 bp were selected and then purified with QIAquick Gel Extraction kit (Qiagen). The DNA fragments were enriched through 10 cycles of PCR using JumpStart Taq DNA Polymerase (Sigma-Aldrich Co., Steinheim am Albuch, Germany), followed by purification using QIAquick Gel Extraction kit (Qiagen) to ensure a fragment size range of

320bp-420bp. Library quality was monitored using an Agilent 2100 BioAnalyzer (Agilent Technologies) and its viable sequencing fragments (molecules carrying adapters at both extremities) was quantified by quantitative PCR using the Library Quantification Kit from KAPA Biosystem. The libraries were amplified using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA USA) on cBot system (Illumina) for cluster generation on the flowcell. Then the amplified flowcell was paired-end sequenced (two reads of 151 bp each) using TruSeq SBS KIT-HS V3 (Illumina) on an Illumina HiSeq X Ten sequencer (Illumina) yielding an average depth of 300 million reads and 48 billion bp of sequence (9X coverage) per sample.

5.2.6. Bioinformatic data analysis

Data filtering was conducted to improve the quality of the DNA reads using an in house BGI pipeline in R software (Version 3.6.1). This included removing adaptor sequences, sequence contaminants and low-quality reads from the raw reads. Low-quality sequences include two types, and any read meeting one or both of the conditions were removed; 1) reads in which unknown bases are > 10%, and 2) reads consisting of 10% or more bases with a Phred quality score of less than 20. The clean reads were then mapped to the *Bos taurus* genome sequence assembly (UMD 3.1.1 with NCBI Accession #: GCF_000003055.6) using the BSMAP program [311]. Duplicate reads were removed, then the mapping results were merged according to UMD 3.1.1 library.

Methyl-cytosine (mC) identification was performed according to the method and correction algorithm of Lister et al. (2009). The methylation level of a cytosine was determined by dividing the number of reads covering each mC by the total reads covering that cytosine [312], which also equals the mC/C ratio at each reference cytosine [313]. The formula is shown below:

$$Rm_{\mathrm{average}} = \frac{\mathrm{Nm_{all}}}{\mathrm{Nm_{all}} + \mathrm{Nnm_{all}}} * 100\%$$

where Nm represents the reads number of mC, while Nnm represents the reads number of non-methylation reads and Rm stands for the mC/C ratio.

The number and proportion of each mC type in the whole genome were classified into three types including CG (or CpG), CHG (H = A, C or T), and CHH, and were derived from the UMD3.1.1 assembly. The methylation levels of different mCs were also mapped to transcriptional units on the bovine genome, specifically denoting methylation levels with respect to 5′-UTR, first exon, first intron, internal exon, internal intron, last exon, and 3′-UTR.

DMRs between Ndiet and Ldiet and between HRFI and LRFI groups were identified using the *de-novo* annotation mode of the Metilene software [314]. Here, a fast circular binary segmentation approach on the mean difference signal of both groups was used [315, 316]. After additional filter steps were applied, potential DMRs were tested using a two-dimensional Kolmogorov-Smirnov-Test (KS-test) [317]. The significance of the DMRs was finally assessed using the Mann-Whitney-U (MWU) test. Potential DMRs had an MWU q-value < 0.05 and minimum mean methylation difference ≥ 0.1.

The subset of filtered significant DMRs identified via Metilene, originally mapped to UMD3.1.1 co-ordinates, were then converted to the corresponding coordinates on the new bovine genome assembly, ARS-UCD1.2 (NCBI Accession: GCF_002263795.1), using the liftOver function within rtracklayer (version 1.44.4) package [318] of R version 3.6.1, and the chain file Tau8ToBosTau9.over.chain, downloaded from the University of California Santa Cruise (UCSC) genome browser's FTP server (http://hgdownload.cse.ucsc.edu/goldenPath/bosTau8/liftOver/). DMRs that did not successfully lift over to the new genome assembly (49 from the total of 673 for Ndiet vs. Ldiet, and 351 from the total of 1666 for HRFI vs. LRFI) because the region was not represented on the ARS-UCD1.2 were discarded from further analysis. As well, 13 prenatal diet-

related DMRs from the UMD assembly were split in the ARS assembly, and for RFI, 60 DMRs were split when transferring RFI-related DMR coordinates from UMD to ARS assemblies. This resulted in the identification of 652 and 1400 DMRs or DMR fragments that were successfully transferred from UMD3.1.1 to ARS-UCD1.2, for the prenatal maternal diet and RFI analyses, respectively. To prepare for gene annotation of the DMRs that successfully were lifted over to the new assembly, the gene transfer format (GTF) file corresponding to the ARS-UCD1.2 assembly was retrieved from the FTP site using the latest release of the Ensembl database (Release 98; September 2019: ftp://ftp.ensembl.org/pub/release-98/gtf/bos_taurus/Bos_taurus.ARS-UCD1.2.98.gtf.gz). The ARS-UCD1.2 assembly GTF file was processed to keep only those columns that are required for DMR annotation in our analysis, after which all duplicated rows were removed. Then we determined the genomic features from the processed GTF file that overlapped with the DMRs successfully lifted over to ARS-UCD1.2. This resulted in the identification of 403 unique DMRs that overlapped with 352 genes for the prenatal diet analysis, and 836 unique DMRs that overlapped with 656 genes for the parental RFI analysis.

5.2.7. Functional enrichment analysis

Fold change and then log2 fold change of the DMR-associated genes for both prenatal diet and parental RFI was determined by dividing the mean methylation of HRFI and Ndiet by the mean methylation of LRFI and Ldiet, respectively. Lists of DMR-associated genes (gene symbols or Ensembl IDs) which consisted of the 352 genes identified for prenatal diet, and 656 identified for divergent parental RFI, along with log2 fold change for each DMR, were imported into the Ingenuity Pathway Analysis (IPA, Ingenuity Systems, Redwood City, CA www.ingenuity.com) software. This software was used to perform gene pathway analysis, gene functional analysis and gene network generation to identify the most significantly perturbed biological pathways affected

by either parental RFI or prenatal maternal diet. The core analysis of IPA was performed to map DMR-associated genes to the Ingenuity Knowledge Base (IKB). The IKB contains large numbers of individually modelled relationships between gene objects (e.g. genes, mRNAs and proteins) to test for significantly over-represented biological networks and pathways using a Fisher's exact test with a cut-off p-value < 0.05. All DMR-associated genes that were associated with a biological term in the IKB were used in the analysis. The significance of the association between the focus genes and the canonical pathway is measured using Fisher's exact test, which is used to calculate a p-value to determine the probability that the network eligible genes that are part of a network appeared there by just random chance. The score is the negative exponent of the right-tailed Fisher's exact test result. For instance, a score of 3 means that probability that the network eligible genes that are part of a network were found there by chance alone is one in 1000. In other words, there is a positive relationship between the number of network eligible genes in a network and score. This means that the higher the number of network eligible genes found in a network, the greater the score (lower the Fisher's exact p-value) that network will have.

5.3. RESULTS

WGBS resulted in a considerable amount of sequence data, ranging from 237,258,977 to 337,297,715 uniquely mapped reads amongst the individual sperm samples, with a mapping rate of 73.6-90.2% (Table 5.1). Bisulfite conversion rates were estimated by inclusion of unmethylated lambda DNA controls along with each sperm sample (Bisulfite Conversion Rate = 1 - methylation rate of lambda DNA). These quality control parameters (Table 5.1) support our ability to faithfully capture patterns of genomic DNA methylation in all sperm samples.

The number and proportion of each mC type (CG, CHG, or CHH) in the whole genome is reported in Table 5.2. Consistent with current patterns of DNA methylation in cattle sperm,

genome wide, CG dinucleotides were preferentially methylated [319]. Across the whole bovine genome, the proportion mCG as compared with mCHG and mCHH was between 95.4 and 97.2% across all samples (Table 5.2). Consistent in all sperm samples and within the transcriptional region, the mean methylation level of CG was highest in the internal intron regions, then in first intron, internal exon, 3′-UTR (downstream), 5′-UTR (upstream), last exon, first exon, and transcriptional start sites (TSS) regions, respectively (Figure 5.1). In other words, the mean methylation level of CG in sperm was highest in the gene body region (including introns and exons), and then in 3′-UTR (downstream), 5′-UTR (upstream), and TSS regions respectively, in agreement with the results of Liu et al. (2019). In terms of the two non-CG methylation types, CHG and CHH, their proportions across the whole bovine genome were 0.8-1.3% and 2-3.3%, respectively. In general, CG methylation is found in both coding genes and in repetitive elements (i.e. short interspersed nuclear elements (SINEs) and long terminal repeats (LTRs)), and is involved in gene expression regulation. On the other hand, the two non-CG methylation types, CHG and CHH, are mostly absent within genes and are mainly found in intergenic regions and repeat-rich regions of the genome [313, 320].

After DMRs were transferred from UMD3.1.1 to the ARS-UCD1.2 genome assembly, 652 DMRs or DMR fragments remained in the comparison between the Ndiet and Ldiet samples. Of these DMRs, 145 regions were hypermethylated and 507 regions were hypomethylated in the Ndiet group. A total of 352 unique genes that overlap with DMRs were imported into IPA. The summary of the IPA core analysis biological functions related to prenatal diet DMR-associated genes is presented in Table 5.3. The significant (p-value < 0.05) functional subcategories of major interest are those associated with Cellular Function and Maintenance, Cell Morphology, as well as Cellular Assembly and Organization in network 1; Organismal Injury and Abnormalities,

Reproductive System Disease, as well as Cancer in network 2; and Connective Tissue Development and Function, Connective Tissue Disorders, as well as Organismal Injury and Abnormalities in network 3. These three gene interaction networks of DMRs obtained from comparisons between the Ndiet and Ldiet bulls are presented in Figures 5.2, 5.3, and 5.4, respectively.

After DMRs were transferred from UMD3.1.1 to the ARS-UCD1.2 genome assembly, 1400 DMRs or DMR fragments remained in the comparison between the HRFI and LRFI samples. Of these DMRs, 798 regions were hypermethylated and 602 regions were hypomethylated in the HRFI group. A total of 656 unique genes that overlap with DMRs were imported into IPA. The summary of the IPA biological functions related to RFI DMR-associated genes is presented in Table 5.4. The significant (p-value < 0.05) subcategories of major interest were those associated with: Embryonic Development, Organ Development, as well as Infectious Diseases in network 1; DNA Replication, Recombination, and Repair, Cell Cycle, as well as Cellular Assembly and Organization in network 2; Embryonic Development, Nervous System Development and Function, as well as Cellular Development in network 3; and Lipid Metabolism, Molecular Transport, as well as Small Molecule Biochemistry in network 4. These four gene interaction networks of DMRs obtained from the comparison between HRFI and LRFI bulls are presented in Figures 5.5, 5.6, 5.7, and 5.8, respectively.

5.4. DISCUSSION

Many of the impacts of maternal dietary perturbations on offspring traits, such as postnatal growth and development, fertility, and general health have been established [3, 4, 85, 305]. Additionally, the relationship between RFI and important traits such as metabolism, fertility, and carcass quality have been well documented [19, 27, 28, 30, 43, 44, 73-75, 321]. However, little is known about

the molecular mechanisms involved in regulating phenotypic responses in the offspring due to differences in prenatal maternal diet and the genetic potential for RFI. It is very plausible that DNA methylation plays a role in both spheres.

The current study reports global bovine sperm methylation profiling using the WGBS to identify DMRs related to the genetic potential for feed efficiency and exposure to prenatal diet in beef bulls. We hypothesized this was important to investigate as it could shed some light on how both genetic (genetic potential for RFI) and environmental (prenatal nutrition) messages are manifested in sperm DNA methylation levels, and potentially passed on to the next generation. We detected 652 (corresponding to 352 unique DMR-associated genes) and 1400 (corresponding to 656 unique DMR-associated genes) DMR and DMR fragments, between normal and low early prenatal nutrition groups, as well as between high and low genetic potential for RFI, respectively, with a q-value < 0.05 in the sperm epigenome. Within the prenatal nutrition comparison, 507 DMRs had enriched methylation levels in Ldiet bulls as compared to 145 DMRs that had enriched methylation levels in the Ndiet bulls. Therefore 77.8% of identified DMRs are more highly methylated in Ldiet bull semen. This is a striking dissimilarity in the direction of methylation differences between Ldiet and Ndiet groups, especially when compared to the RFI genetic potential analysis where 57% of DMRs had enriched methylation levels in the HRFI parental group bulls, as compared with 43% in the LRFI parental group. The greater prevalence of more highly methylated DMRs in sperm of the bulls exposed to undernutrition in utero is in contrast to the study by Radford et al. [322], which saw exactly the opposite. They found that DMRs detected in sperm obtained from adult mice had a higher prevalence of hypo-methylation in individuals born from to undernourished dams. Although, they speculated the DMRs they identified belong to a group that are late to re-methylate in primordial germ cells (PGCs) as fetal development

progresses. Therefore, the profile of DMRs identified by the current study, where the prenatal dietary treatment ended at 150 days of gestation (total gestational length ~285), and the DMRs identified in the mouse study, where undernourishment in utero went from 12.5-18.5 days of gestation (normal gestational length ~19-21 days), which also created distinct differences in birthweight of the offspring [323], might be a different group of DMRs and/or behave quite differently due to timing of prenatal diet/nourishment treat with respect to PCG development timelines. Other bovine sperm methylation studies have seen differences in the number of DMRs that were hypo- versus hyper-methylated, but these comparisons have been mainly performed between high and low fertility sires [324, 325]. Toschi et al. [326] have reported adult sperm DMRs associated with peri-conceptional undernutrition in sheep, but did not indicate any distinct hypo or hyper-methylation patterns within those DMRs. The shear number of DMRs with enriched methylation in Ldiet bulls warrants further investigation.

To understand the possible relevance of the DMRs we have identified, and their potentially associated genes, biological relationships between DMR-associated genes were investigated via IPA. Even though IPA is designed for transcriptomic and proteomic data in order to identify potential cause and effect relationships stemming from differences in gene expression, we used it to identify those biological pathways that might be affected by our two treatments. As a result, three and four statistically significant gene networks were identified for prenatal diet and parental RFI, respectively, as fitting this set of DMR-associated genes. In the following two sections we discuss the possible relevance of these networks, and how they result from or can be related to biological functions associated with: 1) prenatal diet or 2) parental RFI.

5.4.1. DMR-associated gene networks related to prenatal diet

The first network identified for prenatal diet contained DMR-associated with genes relevant to major cellular functions including: a) Cellular Function and Maintenance; b) Cell Morphology; and c) Cellular Assembly and Organization. This network includes 30 DMR-associated genes, of which nine were hypo-methylated while 21 were hyper-methylated in the sperm of Ldiet bulls (Table 5.3 and Figure 5.2). Since genes in this network are connected to the *RAB35* [RAB35, member RAS oncogene family], and *NOTCH3* [notch receptor 3] genes, as well as the AKT [AKT serine/threonine kinase or protein kinase B] signalling pathway, this suggests that this network is primarily associated with cell growth, survival, and proliferation [327-330]. Notch signalling has also been shown to be important for spermatogenesis [331], and the PI3K-AKT pathway important for sperm motility [332]. Alteration of these pathways could possibly affect fertility of the bulls and their sperm, although no obvious differences in sperm fertility parameters between N- and Ldiet bulls were detected [73]. de Castro Barbosa et al. [333] reported that a high fat paternal diet reprogrammed the DNA methylation profile of murine sperm on 92 genes of which some were related to cellular localization, transport, and metabolic processes, somewhat similar to the current study, although none of their DMR-associated genes match those in this first network. The study of Toschi et al. [326] mentioned earlier that would be the most similar to the current study in the animal type used (sheep) and diet treatments applied (peri-conceptual maternal undernutrition), did not identify any similar overrepresented gene ontology groups or DMR-associated genes as was involved in the a) Cellular Function and Maintenance; b) Cell Morphology; and c) Cellular Assembly and Organization network identified by IPA. Although, we have found evidence of differential methylation of genes from this network in sperm from other experimental treatments [334, 335]. Amongst men undergoing IVF, DNA methylation of several genes associated with cell

morphology, cellular assembly and organization, and cellular function and maintenance were detected, plus many were similar to this study [334]. This adds evidence that the methylation status of the DMR-associated genes we have detected from our prenatal diet treatments may be sensitive to environmental factors.

The second network identified for prenatal diet contained DMR-associated with genes relevant to major immune functions including: a) Organismal Injury and Abnormalities; b) Reproductive System Disease; and c) Cancer. This network has 30 DMR-associated genes, of which 23 were hypomethylated and seven were hypermethylated in the sperm of Ndiet bulls (Table 5.3 and Figure 5.3). Since genes in this network are connected to the NF-kB (nuclear factor kappa-light-chain-enhancer of activated B cell) signalling pathway [318, 336] and its associated genes including *INF2* [inverted formin 2] [337], *CACTIN* [cactin, spliceosome C complex subunit] [338], *LRRFIP1* [LRR binding FLII interacting protein 1] [339], the imprinted *TRAPPC9* [trafficking protein particle complex 9] [340], and *ANKRD11* [ankyrin repeat domain 11] [341], this suggests that this network is primarily associated with immune response. NF-kB signalling has also been shown to be associated with pathogenesis of male infertility, with patients having poor sperm concentration had significantly lower levels of NF-kB as compared to control group with normal sperm concentration [342]. *LRRFIP1* mediates the regulation of transcription and translation in the later stage of spermatogenesis [343] and *ANKRD11* is known to be associated with male fertility [344]. Alteration in methylation level of DMR-associated genes in this network could possibly affect fertility of the bulls and their sperm, although no obvious differences in sperm fertility parameters between N- and Ldiet bulls were detected [73]. Nevin and Carroll [345] reported that environmental factors (i.e., diet, smoking and exposure to chemical compounds), testicular injury, sexually transmitted disease and cancer can cause aberrant DNA methylation which impairs sperm

motility, morphology as well as vitality, thereby causing male infertility. Therefore, DNA methylation plays an important role in male fertility. By comparing the sperm methylome of fertile and subfertile Buffalo bulls using a custom-designed microarray, Verma et al. [346] found 151 DMR-associated genes, of which 13 were associated with sperm functions and embryogenesis. Similarly, Camprubí et al. [347] found 696 DMRs associated with 501 genes related to spermatogenesis in the sperm of fertile and infertile humans using the Illumina 450 k array. The study of Toschi et al. [326], which compared the sperm methylome of ram offspring exposed to prenatal undernourishment with a control group, identified one DMR-associated gene, *ST18* [suppression of tumorigenicity 18, zinc finger transcription factor], which also was found in this second network. This supports the fact that different nutritional conditions experienced in utero may alter the methylation status of the sperm genes related to pathogenesis of male infertility, identified in our study.

The third network identified for prenatal diet contained DMR-associated with genes relevant to major cellular functions including: a) Connective Tissue Development and Function; b) Connective Tissue Disorders; and c) Organismal Injury and Abnormalities. This network has 23 DMR-associated genes, of which seven were hypermethylated, while 16 were hypomethylated in the sperm of Ndiet bulls (Table 5.3 and Figure 5.4). Since genes in this network are connected to the *PARP*s [poly(ADP-ribose) polymerases] transcription factor [348], *RAB20* [RAB20, member RAS oncogene family] [327, 328], and *WNT9A* [Wnt family member 9A] [349, 350], this suggests that this network is primarily associated with organismal injury and abnormalities. High *PARP* (a DNA repair enzyme) expression in mature spermatozoa is associated with fertile men, whereas low *PARP* levels in mature sperm is associated with infertile patients, which suggests a role of *PARP* in male infertility [351, 352]. Aberrant methylation of *RAB20* is associated with

infertile human sperm [353]. Regulated Wnt/Beta-Catenin signaling is essential for diverse processes during development and in adult tissue homeostasis and sustains adult spermatogenesis in Mice [354, 355]. Alteration of these pathways could possibly affect fertility of the bulls and their sperm, although no obvious differences in sperm fertility parameters between N- and Ldiet bulls were detected [73]. The study of Toschi et al. [326] did identify two similar DMR-associated genes including *HDAC4* [histone deacetylase 4] and *HNRNPUL1* [heterogeneous nuclear ribonucleoprotein U like 1] which were also identified in this third network. Similar to the second network, this overlap in DMR associated genes supports the fact that different nutritional conditions experienced in utero may alter methylation status of gene networks involved in a) Connective Tissue Development and Function; b) Connective Tissue Disorders; and c) Organismal Injury and Abnormalities.

5.4.2. DMR-associated gene networks related to parental RFI

The first network identified for RFI contained DMR-associated with genes relevant to major cellular functions including a) Embryonic Development; b) Organ Development; and c) Infectious Diseases. This network includes 27 DMR-associated genes. Seventeen genes were hypermethylated, while 10 were hypomethylated in the sperm of HRFI bulls (Table 5.4 and Figure 5.5). Since genes in this network are connected to regulation of *ERK1/2* [mitogen-activated protein kinase 3/1 or *MAPK3/1*] signalling pathway [356-358], *IGF2R* [279, 359], as well as Rab family of proteins including *RABL6* [RAB, member RAS oncogene family like 6], *RAB7A* [RAB7A, member RAS oncogene family], and *RAB7B* [RAB7B, member RAS oncogene family] [327, 328], this suggests that this network is primarily associated with growth and development. We did not see any difference in overall growth between our HRFI and LRFI bulls, however we did see that LRFI bulls had a smaller scrotal circumference than HRFI bulls [73]. Zhang et al. [360] reported

that LRFI lambs had smaller rumen and longer duodenum (indicating the LRFI lambs had less feed intake and more efficient absorption rate than HRFI lambs), smaller liver, lung and kidney than HRFI lambs. Meyer et al. [361] also reported that LRFI lamb tended to have larger pancreas and spleen than HRFI lambs. Fitzsimons et al. [362] reported lighter weight of empty reticulo-rumen in LRFI compared with HRFI bulls. Basarab et al. [16] also reported lower weight of liver, stomach and intestine in LRFI steers compared to their HRFI counterparts. Additionally, we also seen differences in the expression of some genes associated with growth and development in *Longissimus thoracis* (LT) muscle, semimembranosus (SM) muscle, liver, and/or testes (as reported in chapter 4). The study of Karisa et al. [145] found SNP-associated genes related to regulatory pathways controlling growth and development in divergent RFI beef cattle such as *ERK1/2* reported in this network, although it was not a DMR-associated gene. These findings suggest that the in terms of size/weight of different organs, at least regarding visceral organs, there might be some differences between HRFI and LRFI animals, for which these pathways could play a part of. *ERK* or *MAPK* cascade regulates normal spermatogenesis and germ cell functions [363]. *RABL6*, *RAB7A*, *RAB7B* belong to Rab family of proteins are key regulators of male fertility [364]. Although no obvious differences in sperm fertility parameters between HRFI and LRFI bulls were detected and only LRFI bulls tended to exhibit lower progressive motility compared with high RFI bulls [73], alteration of these pathways may affect fertility of the bulls and their sperm. Indeed, there is only one study in cattle that compared global methylation pattern of HRFI and LRFI in liver [191] but not in sperm. In this study which was conducted by Rocha et al. [191], 1493 differentially methylated cytosines (DMCs) and 279 differentially methylated regions (DMRs) were identified in the hepatic tissue of Nelore cattle exhibiting extremes in RFI, although the

association between these DMCs and DMRs with genes or biological pathways was not investigated.

The second network found for RFI had DMRs which were particularly associated with genes related to a) DNA Replication, DNA Repair, and RNA Processing; b) Cell Cycle; and c) Cellular Assembly and Organization. It consists of 26 DMR-associated genes, of which 21 were hypermethylated, while five were hypomethylated in the sperm of HRFI bulls (Table 5.4 and Figure 5.6). The term cellular assembly and organization refers to nuclear assembly and the organization of genes and chromosomes which occurs at three levels including [365]: 1) organization of nuclear processes themselves, including DNA replication, DNA repair, transcription, and RNA processing; 2) organization of chromatin; and 3) arrangement of chromosomes and genes within the nuclear space. Since genes in this network are connected to DMR-associated genes including *SMARCB1* [SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1] [366] and *POLR2A* [RNA polymerase II subunit A] [367], this suggests that this network is primarily associated with regulation of gene expression, DNA replication, and DNA repair. Ramayo-Caldas et al. [368] reported that *POLR2A* and *CHD1* [chromodomain helicase DNA binding protein 1] genes were differentially co-expressed in divergent RFI pigs, which could substantiate our identification of these genes being RFI DMR-associated genes. With respect to potential fertility differences, aberrant methylation of *SMARCB1* is associated with infertile human sperm [353], and alterations in the transcription factor binding site of *POLR2A* is associated with infertile human sperm [369, 370]. The LRFI bulls in this study tended to exhibit lower progressive motility compared with high RFI bulls [73]. Therefore, alteration of these pathways may affect fertility of the bulls and their sperm.

Similar to the first network with respect to RFI, the third network found for RFI had DMRs which were particularly associated with genes related to a) Embryonic Development; b) Nervous System Development and Function; and c) Cellular Development. It consists of 21 DMR-associated genes, of which 15 were hypermethylated, whereas six were hypomethylated in the sperm of HRFI bulls (Table 5.4 and Figure 5.7). Since the genes in this network are connected to *PI3K* [phosphatidylinositol 3-kinase] (Hemmings and Restuccia, 2012) and *CREBBP* [CREB binding protein] (Tang et al., 2016), it suggests that this network is primarily associated with metabolic processes involved in growth and development. As mentioned for the first network, there are differences in size/weight of organs in HRFI vs. LRFI animals [16, 360-362]. We have also seen smaller scrotal circumference in LRFI bulls [73] and different expression of genes associated with growth and development in *Longissimus thoracis* (LT) muscle, semimembranosus (SM) muscle, liver, and/or testes (as reported in chapter 4). Karisa et al. [145] reported SNP-associated genes related to regulatory pathways controlling growth and development in divergent RFI beef cattle such as *JAK* [Janus kinase] and *STAT5* [signal transducer and activator of transcription 5] reported in this network, although they were not DMR-associated genes. Both *PI3K* [371] and *CREBBP* [353] also regulate normal spermatogenesis and are important regulators of male fertility. As mentioned earlier, LRFI bulls in this study showed a tendency of lower progressive motility compared with high RFI bulls [73]. Hence, alteration of these pathways may also affect fertility of the bulls and their sperm.

The fourth network found for RFI had DMRs which were particularly associated with genes related to a) Lipid Metabolism; b) Molecular Transport; and c) Small Molecule Biochemistry. This network consists of 14 DMR-associated genes, of which three were hypo- and 11 were hyper-methylated in the sperm of HRFI bulls (Table 5.4 and Figure 5.8). Since genes in

this network (i.e., *PNPLA4* [patatin like phospholipase domain containing 4], *CPT1C* [carnitine palmitoyltransferase 1C], *ABCA3* [ATP binding cassette subfamily A member 3], and *CLYBL* [citrate lyase beta like]) are connected to the production of beta-estradiol and triacylglycerol, this suggests that this network is primarily associated with lipid or steroid metabolism. Several studies looking at ultrasound scans of subcutaneous fat in cattle suggested that selection for LRFI cattle resulted in progeny steers having more whole-body chemical protein and less whole-body chemical fat than the progeny of HRFI parents [22, 30]. In addition, study of Weber et al. [372] showed down-regulation of regulatory pathways controlling fat deposition in LRFI cattle such as lipid metabolism, similar to our study. As well, *PNPLA4*, *CPT1C*, *ABCA3*, and *CLYBL* regulate normal spermatogenesis and are important regulators of male fertility [353]. Hence, changes in these pathways may also affect fertility of the bulls and their sperm.

5.5. CONCLUSIONS

In this study, we performed global DNA methylation profiling to elucidate the underlying epigenetic mechanisms of the effects of prenatal maternal diet or selection for parental RFI in cattle offspring at the level of sperm DNA. Comparison between Ndiet and Ldiet groups identified 652 DMRs (corresponds to 352 unique DMR-associated genes) in the sperm of bull offspring. In addition, comparison between HRFI and LRFI parental groups identified 1400 DMRs (corresponds to 656 unique DMR-associated genes) in the sperm of the bull offspring. The Ndiet versus Ldiet comparison displayed a striking difference in the number of hyper- and hypo-methylated DMRs, with 77.8% hyper-methylated in Ldiet bulls. This is in contrast with the RFI comparison where HRFI sperm displayed 57% of the hyper-methlyated DMRs. To understand the possible functions of the DMRs and their potentially associated genes, biological relationships were determined by pathway enrichment analysis via IPA. As a result, three networks associated

with “cell survival and growth”, “disease or abnormalities”, and “connective tissue development” were identified as being different when comparing the Ndiet group to the Ldiet group. Many of the genes in these networks are important for normal spermatogenesis and male fertility. Similar pathway analysis for the HRFI and LRFI bulls showed difference in four networks involved in “embryonic development”, “DNA replication, DNA repair, and RNA processing”, “growth control and homeostasis”, as well as “lipid metabolism”. These data suggest that both prenatal under-nutrition and parental RFI selection can alter the epigenome of sperm cells in progeny bulls.

Table 5. 1 Number of unique mapped reads, mapping rate, and bisulfite conversion rate amongst the bovine sperm samples obtained by the whole genome bisulfite sequencing (WGBS) technique

Sperm ID	Mapped Reads [a]	Mapping Rate [b] (%)	Bisulfite Conversion Rate [c] (%)
Sperm701	296,499,021	88.9	99.7
Sperm703	264,469,768	88.2	99.7
Sperm707	291,661,108	87.7	99.6
Sperm709	274,850,036	88.8	99.7
Sperm711	320,324,688	85.0	99.6
Sperm713	302,505,051	85.6	99.6
Sperm715	270,111,566	90.2	99.6
Sperm717	243,556,545	83.5	99.7
Sperm719	268,759,310	81.4	99.6
Sperm721	248,850,976	89.0	99.7
Sperm723	257,337,584	83.7	99.6
Sperm725	337,297,715	88.8	99.6
Sperm727	280,060,696	83.9	99.6
Sperm729	243,224,291	85.0	99.6
Sperm735	295,178,407	87.2	99.6
Sperm737	281,289,208	83.5	99.6
Sperm739	239,821,806	88.2	99.6
Sperm741	290,135,716	81.5	99.7
Sperm749	282,042,403	73.6	99.7
Sperm751	277,129,595	87.9	99.7
Sperm753	245,355,343	82.6	99.6
Sperm755	316,136,068	85.9	99.7
Sperm795	328,046,844	88.8	99.7
Sperm817	237,258,977	81.9	99.6
Sperm823	264,430,244	83.6	99.7

[a] Mapped reads: total read numbers that can be mapped to the reference genome; [b] Mapping rate: the proportion of reads that can be mapped to reference genome out of the total number of reads for each sample; [c] Bisufite conversion rate: the proportion of non-methylated cytosine that has been converted to uracils by bisulfite treatment in total non-methylated cytosine.

Table 5. 2 Number and proportion of mCG, mCHG (H = A, C or T) and mCHH of all methyl-cytosines in each sample

Sample ID		mCG	mCHG	mCHH
Sperm701	mC number	22,639,020	213,186	560,132
	proportion (%) [a]	96.7	0.9	2.4
Sperm703	mC number	19,656,408	165,302	419,442
	proportion (%)	97.1	0.8	2.1
Sperm707	mC number	24,371,007	229,797	635,349
	proportion (%)	96.6	0.9	2.5
Sperm709	mC number	20,093,435	164,484	414,604
	proportion (%)	97.2	0.8	2.0
Sperm711	mC number	25,749,778	266,992	750,733
	proportion (%)	96.2	1.0	2.8
Sperm713	mC number	20,405,393	217,946	602,171
	proportion (%)	96.1	1.0	2.8
Sperm715	mC number	20,575,533	168,214	423,174
	proportion (%)	97.2	0.8	2.0
Sperm717	mC number	20,979,878	167,961	437,408
	proportion (%)	97.2	0.8	2.0
Sperm719	mC number	23,586,378	218,168	597,425
	proportion (%)	96.7	0.9	2.4
Sperm721	mC number	19,340,705	163,231	412,014
	proportion (%)	97.1	0.8	2.1
Sperm723	mC number	22,524,640	194,086	509,822
	proportion (%)	97.0	0.8	2.2
Sperm725	mC number	23,294,852	244,017	670,007
	proportion (%)	96.2	1.0	2.8
Sperm727	mC number	20,067,219	265,412	694,471
	proportion (%)	95.4	1.3	3.3
Sperm729	mC number	19,795,405	187,896	514,616
	proportion (%)	96.6	0.9	2.5
Sperm735	mC number	20,029,505	195,887	488,106
	proportion (%)	96.7	0.9	2.4
Sperm737	mC number	24,760,884	255,042	728,872
	proportion (%)	96.2	1.0	2.8
Sperm739	mC number	19,680,793	168,611	417,309
	proportion (%)	97.1	0.8	2.1
Sperm741	mC number	24,975,324	206,417	577,221
	proportion (%)	97.0	0.8	2.2
Sperm749	mC number	20,357,930	235,597	666,279
	proportion (%)	95.8	1.1	3.1
Sperm751	mC number	21,101,801	189,054	490,934
	proportion (%)	96.9	0.9	2.3
Sperm753	mC number	22,902,760	211,505	589,283
	proportion (%)	96.6	0.9	2.5
Sperm755	mC number	24,670,387	208,577	555,554
	proportion (%)	97.0	0.8	2.2
Sperm795	mC number	25,058,783	283,578	779,983
	proportion (%)	95.9	1.1	3.0
Sperm817	mC number	20,644,327	203,495	581,400
	proportion (%)	96.3	1.0	2.7
Sperm823	mC number	20,495,156	186,390	475,767
	proportion (%)	96.9	0.9	2.2

[a] The sum of proportion of three mC types is 100%.

Table 5. 3 DMR-associated gene networks related to prenatal diet identified by IPA

ID	Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
1	↓ACAD9, Akt, ↑**BAIAP2L1**, ↓BPIFB2, ↓DENND1B, ↑**EVPL**, ↓FBXL17, ↓FLYWCH1, ↓FRMD1, ↑**GEM**, ↓GRB10, ↓INPP4A, ↑**ITPR3**, ↓KIAA1671, Mlc, Mlcp, ↓MPRIP, Myosin, ↓NDUFV1, ↓NISCH, ↑**NOTCH3**, Pkg, ↓PPP1R12B, ↓PRDM16, ↓PRMT8, ↓RAB35, ↓SCD5, ↑**SHANK2**, ↓TBC1D10A, ↑**TFAP4**, ↑**UNC45B**, ↓VTCN1, ↓WIPF2, ↑**ZDHHC8**, ↓ZNF516	55	30	Cellular Function and Maintenance; Cell Morphology; Cellular Assembly and Organization
2	↓ACADVL, ↓AFAP1, ↓AFF1, ↓ANKRD11, c-Src, ↓CACTIN, ↓CEP170B, ↑**CLASRP**, ↓ERC2, ↓FAM83H, ↓FRY, ↓FYCO1, ↓GAK, ↑**GRAMD4**, ↓ICE1, ↑**IFT122**, ↓INF2, ↑**IQSEC1**, ↓LRRFIP1, MAP1LC3, mediator, NFkB (complex), P-TEFb, ↓P2RX3, ↑**PADI2**, ↓PDE4DIP, ↓PKN3, ↓RBM10, ↑**SASH1**, ↓SLCO3A1, ↓ST18, ↓TECPR1, ↓TRAPPC12, ↓TRAPPC9, ↑**ZFP64**	55	30	Organismal Injury and Abnormalities; Reproductive System Disease; Cancer
3	↓ATP11A, ↓BCL2L11, ↑**CAND2**, caspase, cytochrome C, ↓DNAJB6, ↓ELP4, ↓GMDS, ↓HCN1, Hdac, ↓HDAC4,histone deacetylase, Histone h4, ↓HNRNPUL1, Hsp27, Hsp70, Hsp90, ↑**INTS1**, Jnk, ↓KLF6, ↓LRP4, ↑**MAPK8IP1**, ↑**MCPH1**, N-cor, Nucleoporin, ↓PACS2, PARP, ↓RAB20, ↓RCOR1, ↓RGL1, ↑**TBX1**, ↑**TNKS2**, ↓TTLL4, ↓WIPI2, ↑**WNT9A**	37	23	Connective Tissue Development and Function; Connective Tissue Disorders; Organismal Injury and Abnormalities

DMR-associated genes are underlined and hypermethylated DMR-associated genes in Ndiet group are marked in bold.

Table 5. 4 DMR-associated gene networks related to genetic potential for RFI identified by IPA

ID	Molecules in Network	Score	Focus Molecules	Top Diseases and Functions
1	Adaptor protein, Adaptor protein 1, ↓AHDC1, Ap1 gamma, **↑AP1B1**, Ap2 alpha, ↓AP2A2, **↑B3GAT1**, ↓CLTB, **↑COMMD9**, **↑CROCC**, **↑DENND1B**, Dgk, **↑DGKQ**, **↑DGKZ**, ↓DOK5, Dynamin, ↓EPN1, ERK1/2, ↓GAREM1, **↑GTSE1**, **↑HIP1**, **↑IGF2R**, ↓LRRC8E, **↑MGRN1**, RAB7, **↑RAB7A**, **↑RAB7B**, **↑RABL6**, **↑SUSD5**, ↓TAFA4, **↑TBC1D5**, ↓TRIM2, **↑WASHC2A/WASHC2C**, ↓ZYG11B	40	27	Embryonic Development; Organ Development; Infectious Diseases
2	**↑CARS2**, **↑CHD1**, **↑CHD7**, ↓COL6A1, **↑COL6A5**, Cyclin A, Cyclin E, **↑DYNC1H1**, E2f, **↑ESYT2**, ↓GATAD2A, **↑IFT140**, **↑INIP**, Insulin, **↑KAT2B**, **↑LARS2**, ↓LYRM4, **↑MCM4**, **↑MED15**, ↓MED6, mediator, **↑MLLT1**, **↑POLR2A**, **↑RAB36**, Rb, **↑SDHB**, **↑SMARCB1**, **↑SMOC2**, **↑TADA2B**, **↑TFDP1**, TFIIH, thymidine kinase, **↑TLN2**, TRAP/Media, ↓ZNF592	37	26	DNA Replication, DNA Repair, and RNA Processing; Cell Cycle; Cellular Assembly and Organization
3	amylase, ↓ARHGEF10L, Atrial Natriuretic Peptide, **↑BAHCC1**, Cbp/p300, ↓CLP1, **↑CREBBP**, **↑EBF1**, **↑EBF3**, EGLN, **↑EP400**, **↑FOXK1**, ↓GIMAP4, Ifn, IFN Beta, IFN type 1, **↑IFNW1**, ↓IMMP2L, IRF, ↓IRF2, JAK, **↑KIF26A**, ↓LSP1, NCOA, **↑PCGF3**, PEPCK, PI3K (complex), **↑PLEKHF1**, **↑RNF165**, **↑SETD3**, SMAD1/5/9, **↑SPDEF**, STAT5a/b, **↑ZFHX3**, **↑ZNF511**	27	21	Embryonic Development; Nervous System Development and Function; Cellular Development
4	AADAC, **↑ABCA3**, ARID5B, **↑ART5**, beta-estradiol, CCR3, CHRM3, **↑CLYBL**, ↓CPT1C, GALR1, GPR137B, GPR37L1, GPR88, GRM3, HDAC1, **↑HMCN2**, HR, **↑KCNIP1**, ↓LRFN3, **↑NACC2**, **↑NALCN**, **↑PADI4**, **↑PHF21B**, **↑PNPLA4**, PON2, PTCH1, **↑RAPSN**, RNF130, RXFP2, SCTR, SEMA3C, ST6GALNAC3, TNNI2, triacylglycerol, ↓UGT3A2	15	14	Lipid Metabolism; Molecular Transport; Small Molecule Biochemistry

DMR-associated genes are underlined and hypermethylated DMR-associated genes in HRFI group are marked in bold.

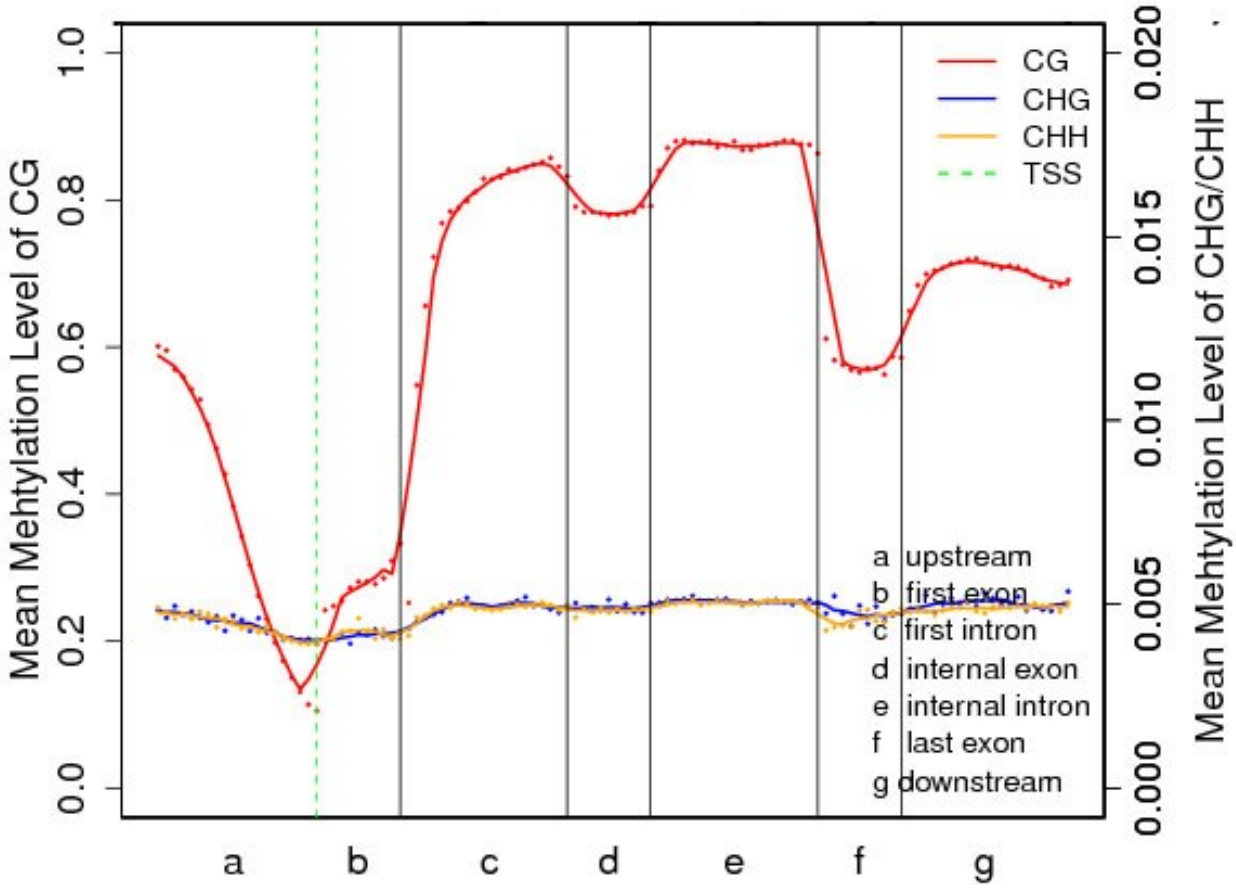

Figure 5. 1 Methylation trends within the transcriptional unit for sperm701. Methylation level for CG (or CpG), CHG (H = A, C or T), and CHH are shown with red, blue, and orange colors, respectively. TSS stands for transcriptional start site and is shown by the dotted green line. Upstream refers to 5′-UTR, whereas downstream refers to 3′-UTR.

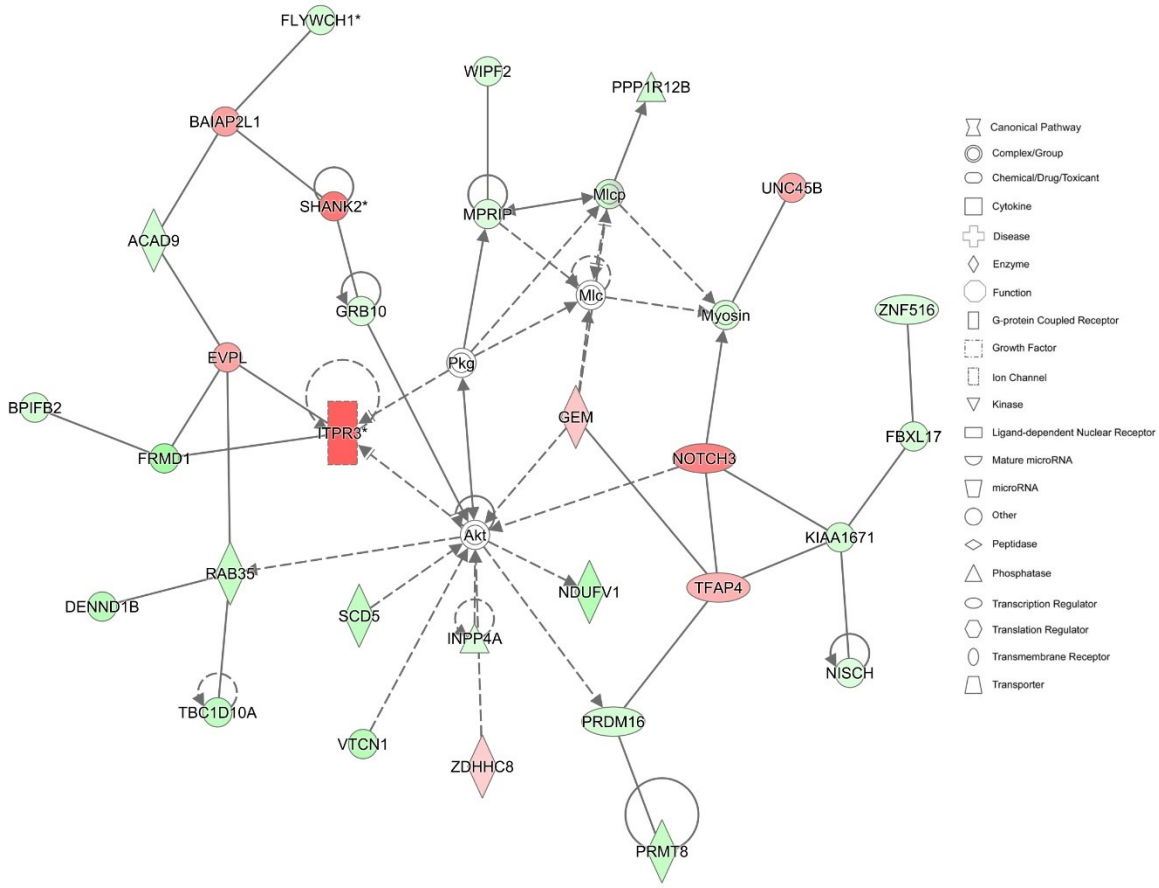

Figure 5. 2 The relationships between DMR-associated gene with respect to prenatal diet in network 1, which have functions in Cellular Function and Maintenance, Cell Morphology, as well as Cellular Assembly and Organization. DMR-associated genes coloured in red are hypermethylated in Ndiet bulls. DMR-associated genes coloured in green are hypermethylated in Ldiet bulls. The colour intensity correlates to the level of methylation.

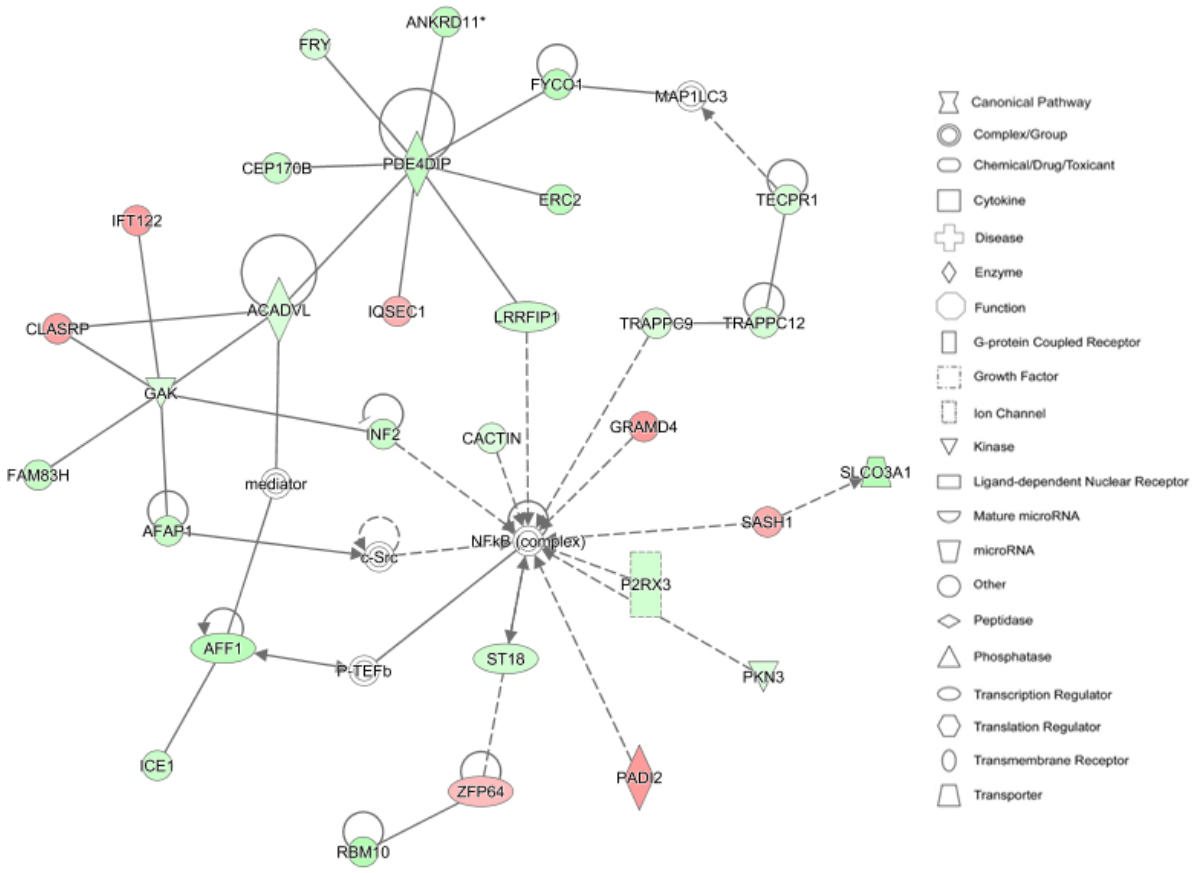

Figure 5. 3 The relationships between DMR-associated gene with respect to prenatal diet in network 2, which have functions in Organismal Injury and Abnormalities, Reproductive System Disease, as well as Cancer. DMR-associated genes coloured in red are hypermethylated in Ndiet bulls. DMR-associated genes coloured in green are hypermethylated in Ldiet bulls. The colour intensity correlates to the level of methylation.

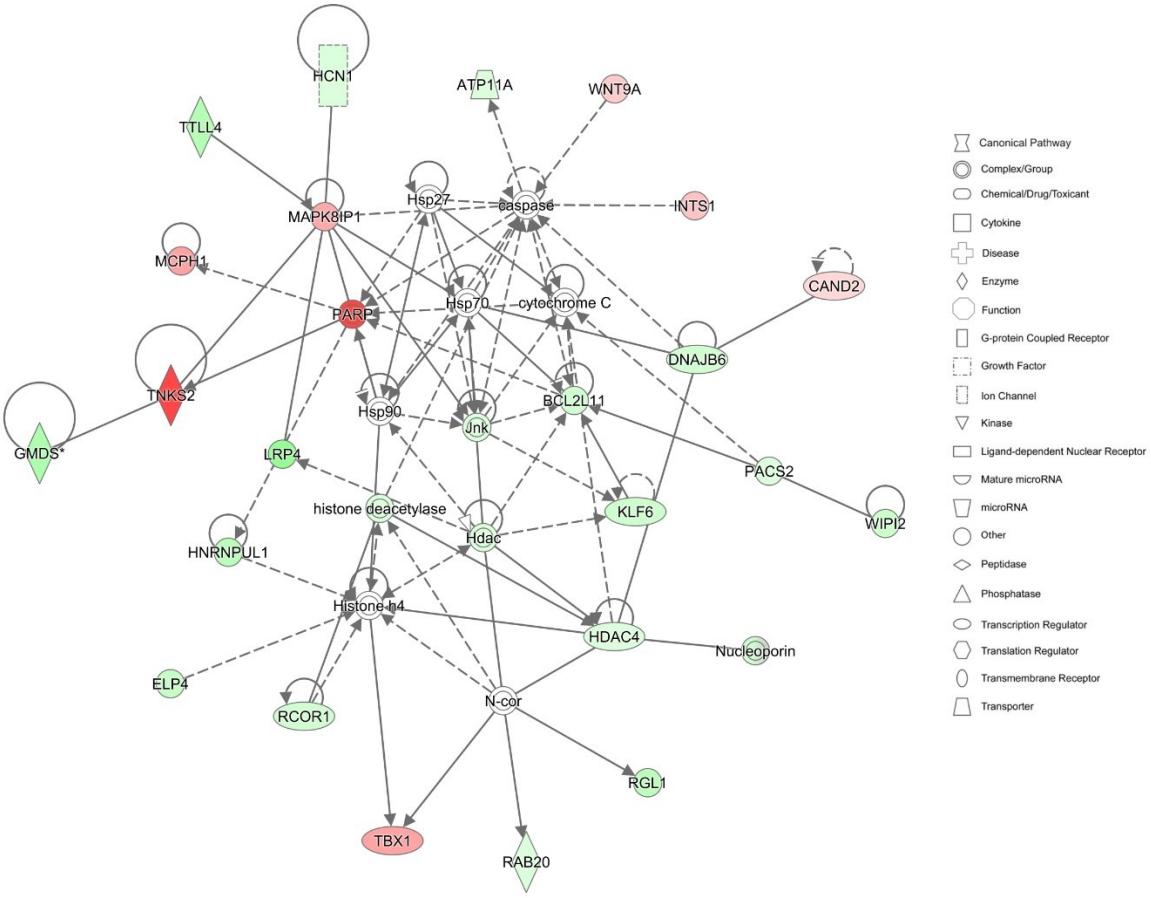

Figure 5. 4 The relationships between DMR-associated gene with respect to prenatal diet in network 3, which have functions in Connective Tissue Development and Function, Connective Tissue Disorders, as well as Organismal Injury and Abnormalities. DMR-associated genes coloured in red are hypermethylated in Ndiet bulls. DMR-associated genes coloured in green are hypermethylated in Ldiet bulls. The colour intensity correlates to the level of methylation.

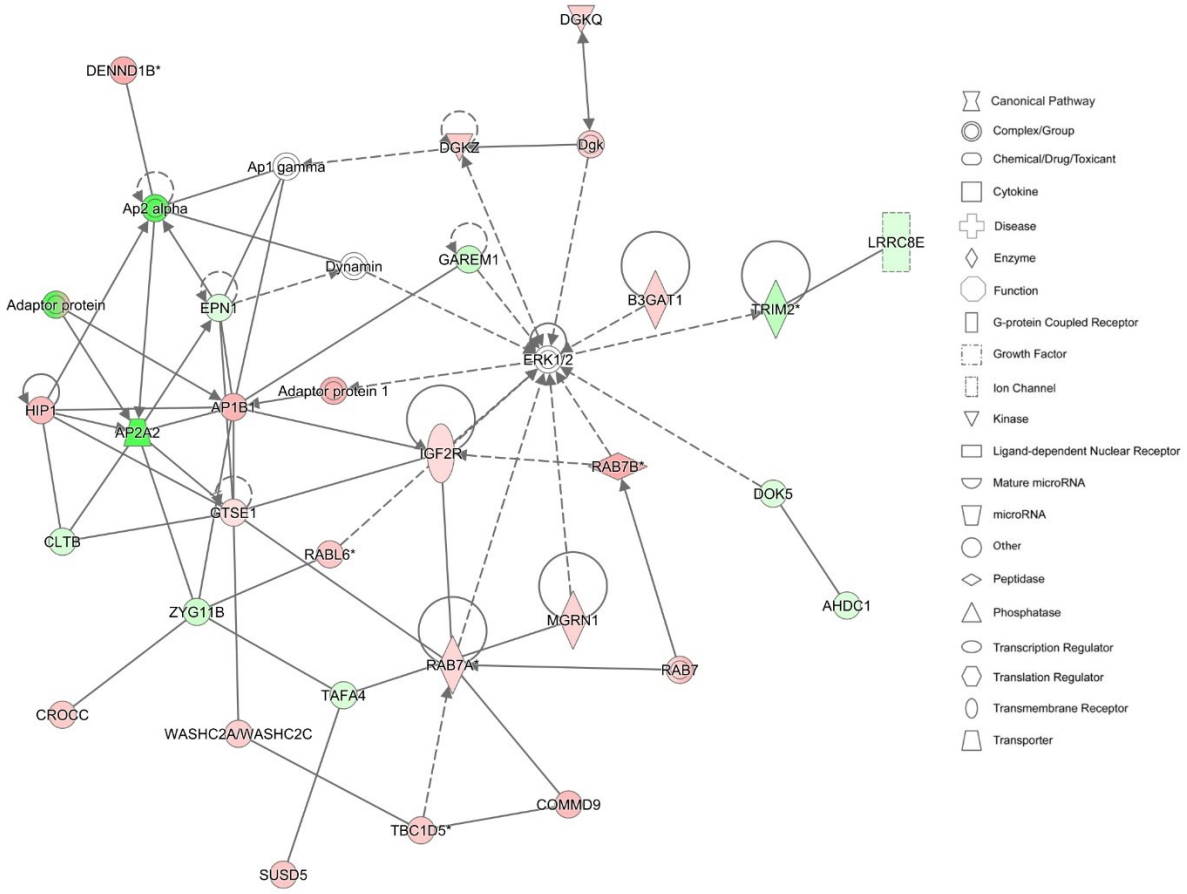

Figure 5. 5 The relationships between DMR-associated gene with respect to genetic potential for RFI in network 1, which have functions in Embryonic Development, Organ Development, Infectious Diseases. DMR-associated genes coloured in red are hypermethylated in HRFI bulls. DMR-associated genes coloured in green are hypermethylated in LRFI bulls. The colour intensity correlates to the level of methylation.

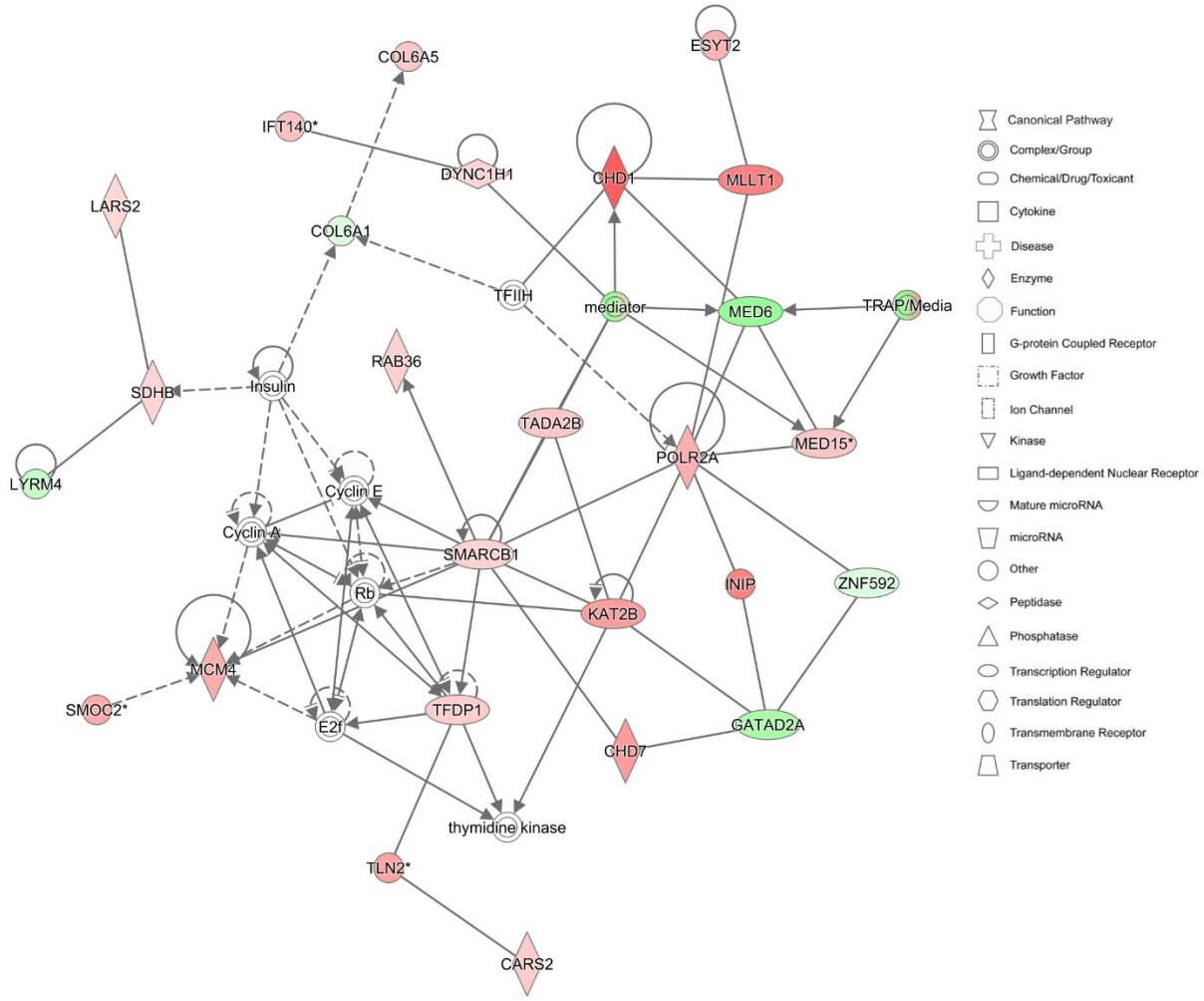

Figure 5. 6 The relationships between DMR-associated gene with respect to genetic potential for RFI in network 2, which have functions in DNA Replication, Recombination, and Repair, Cell Cycle, as well as Cellular Assembly and Organization. DMR-associated genes coloured in red are hypermethylated in HRFI bulls. DMR-associated genes coloured in green are hypermethylated in LRFI bulls. The colour intensity correlates to the level of methylation.

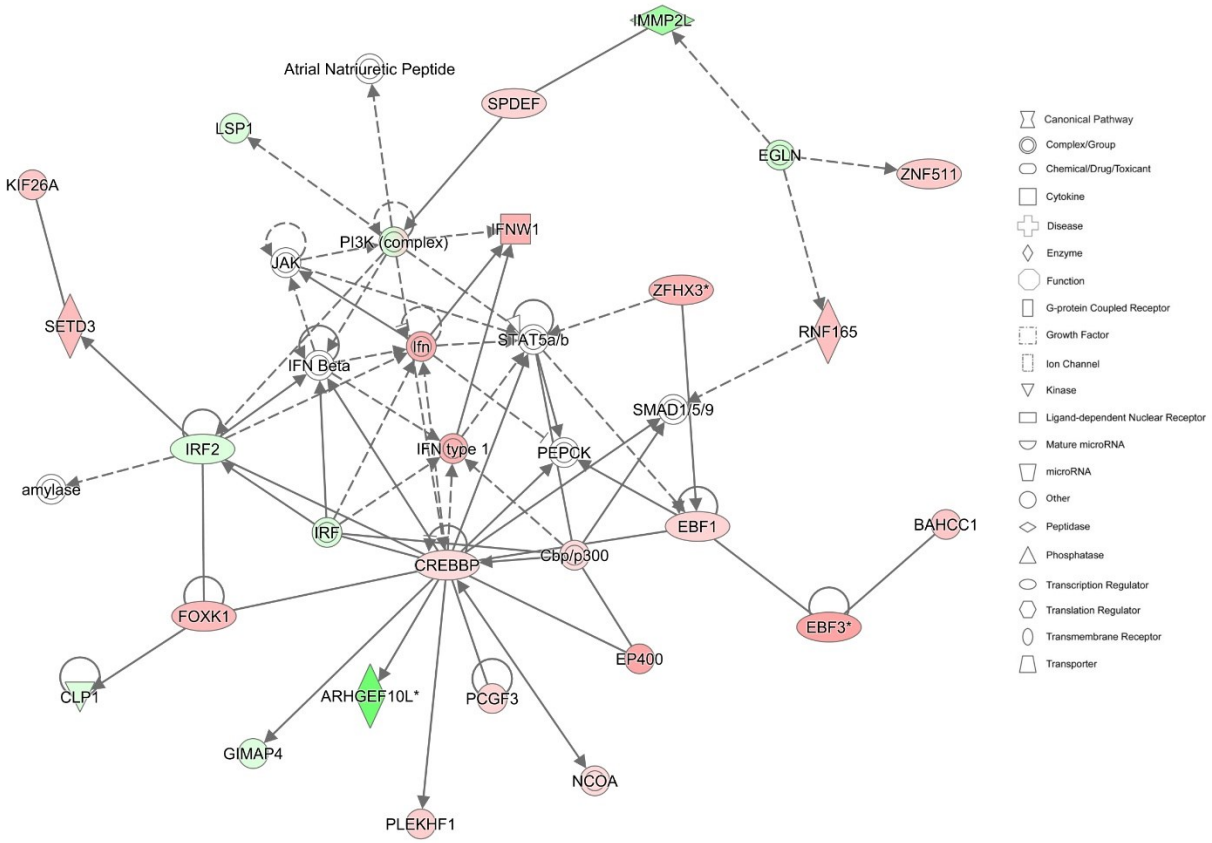

Figure 5. 7 The relationships between DMR-associated gene with respect to genetic potential for RFI in network 3, which have functions in Embryonic Development, Nervous System Development and Function, as well as Cellular Development. DMR-associated genes coloured in red are hypermethylated in HRFI bulls. DMR-associated genes coloured in green are hypermethylated in LRFI bulls. The colour intensity correlates to the level of methylation.

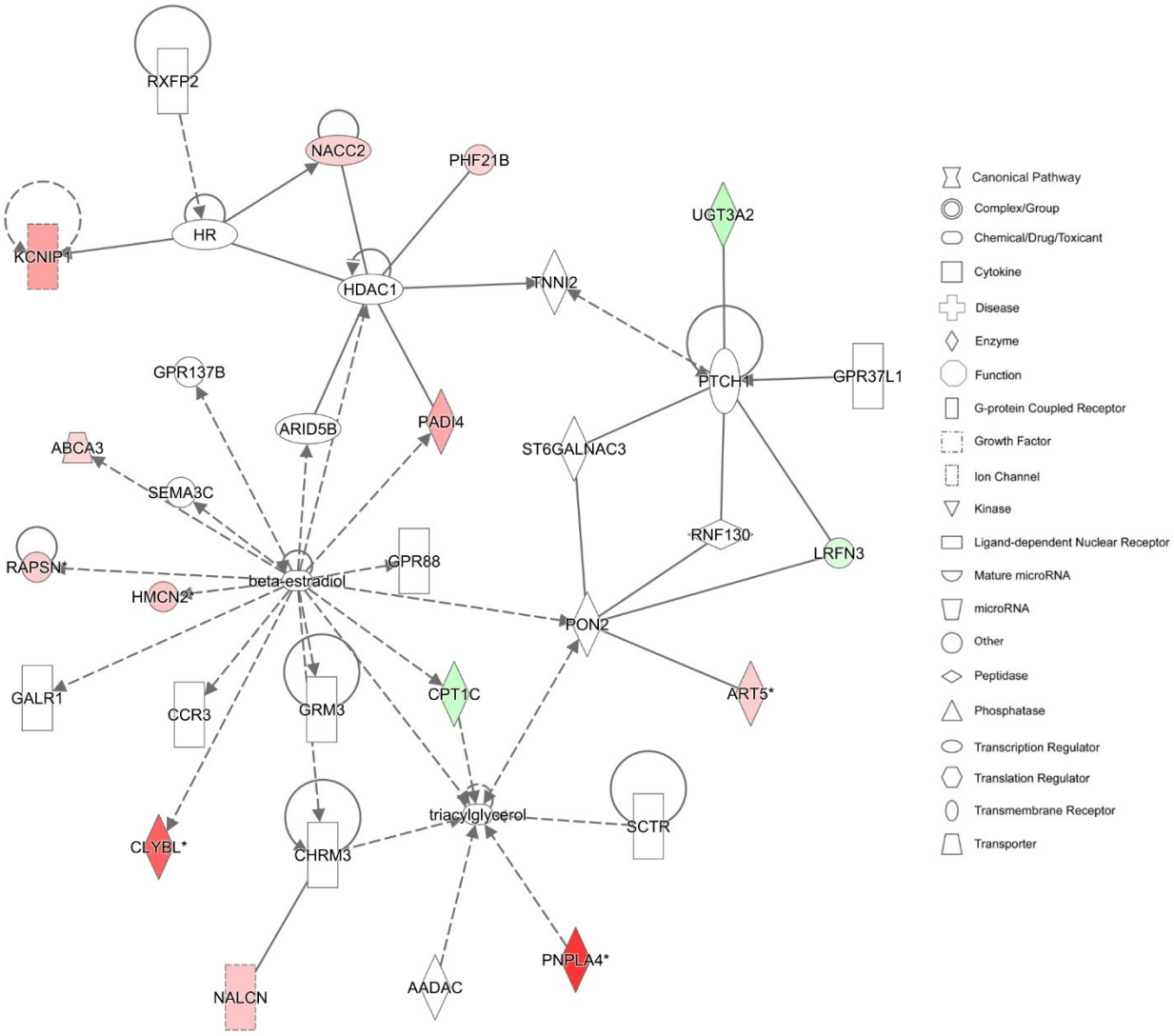

Figure 5. 8 The relationships between DMR-associated gene with respect to genetic potential for RFI in network 4, which have functions in Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry. DMR-associated genes coloured in red are hypermethylated in HRFI bulls. DMR-associated genes coloured in green are hypermethylated in LRFI bulls. The colour intensity correlates to the level of methylation.

CHAPTER 6. GENERAL CONCLUSION

My thesis investigated four different hypotheses: 1) The bovine metabolome can be comprehensively described and annotated using comprehensive, quantitative metabolomics techniques and computer-aided literature mining; 2) Differences between HRFI and LRFI bulls can be detected using metabolomics techniques, such that serum-based metabolite biomarker panels can be developed to distinguish HRFI from LRFI animals; 3) Differences in selection for parental RFI and/or prenatal nutrition can change gene expression patterns in tissues of young Angus bulls leading to changes in immunity and muscle development; 4) Differences in selection for parental RFI and/or prenatal nutrition can change DNA methylation patterns in the sperm of young Angus bulls. These hypotheses were described and tested in chapters 2-5 of this thesis. Here I will summarize the importance of each hypothesis/chapter and their associated results:

Hypothesis 1/Chapter 2. From an animal health perspective, relatively little is known about the typical or healthy ranges of concentrations for many metabolites in bovine biofluids and tissues. In an effort to gain a better understanding of bovine metabolism and a better picture of the chemical composition of bovine milk/meat products, I hypothesized that the bovine metabolome could be comprehensively described and annotated using quantitative metabolomics techniques and computer-aided literature mining. In testing this hypothesis in chapter 2, I described the results of a comprehensive, quantitative metabolomic characterization of six bovine biofluids and tissues, including serum, ruminal fluid, liver, LT muscle, SM muscle, and testes tissues. Using NMR, LC-MS/MS, and ICP-MS, we were able to identify and quantify more than 145 metabolites in each of these biofluids/tissues. Combining these results with previous work done on other bovine biofluids, as well as previously published literature values for other bovine tissues and biofluids, we were able to generate quantitative reference concentration data for 2100 unique metabolites

across five different bovine biofluids and seven different tissues. These experimental data were combined with computer-aided, genome-scale metabolite inference techniques to add another 48,628 unique metabolites that are biochemically expected to be in bovine tissues or biofluids. Altogether, 51,801 unique metabolites were identified in this study. Detailed information for these 51,801 unique metabolites has been placed in a publicly available database called the Bovine Metabolome Database (www.bovinedb.ca). I believe these reference data will be useful for understanding more about bovine biology, for assessing the micronutrients found in bovine tissues as well as for improving the veterinary care of beef and dairy cattle.

Hypothesis 2/Chapter 3. RFI is a feed efficiency measure commonly used in the livestock industry to identify animals that efficiently or inefficiently convert feed into meat or body mass. Selection for LRFI or feed efficient animals is gaining popularity among beef producers due to the fact that LRFI cattle eat less and produce less methane per unit weight gain. Several studies have indicated that LRFI cattle have reduced feed intake by 9-15% at equal weight and average daily gain (ADG) [16, 22-24]. However, some studies have shown that some fertility issues may exist with certain LRFI animals [76, 77, 79]. LRFI cattle also exhibited an improved feed conversion ratio (FCR) by 10-15% at equal weight and ADG, compared to their High-RFI (HRFI) counterparts [16, 22, 24]. In addition, RFI is moderately heritable in cattle, which makes it a good candidate for genetic improvement. RFI is a difficult and time-consuming measure to perform and therefore a simple blood test that could distinguish HRFI from LRFI animals (early on) would potentially benefit beef farmers in terms of optimizing production or selecting which animals to cull or which animals should be bred. Serum and plasma metabolite differences between HRFI vs. LRFI cattle have been reported previously, which suggests that a simple blood test for RFI testing may be possible. However, the results have been variable and have never been replicated nor expressed in

a clear biomarker profile. Therefore in chapter 3, I hypothesized that differences between HRFI and LRFI bulls could be detected using metabolomics techniques, such that serum-based metabolite biomarker panels could be developed to distinguish HRFI from LRFI animals. In testing this hypothesis for chapter 3, I employed quantitative NMR spectroscopy, LC-MS/MS, and ICP-MS techniques and combined the results with logistic regression to identify two serum biomarker panels for RFI prediction. These include an NMR-based two-metabolite model included formate and leucine, an MS-based two-metabolite model included C4 (butyrylcarnitine) and LysoPC(28:0). These serum biomarkers have high sensitivity and specificity (AUROC > 0.85), for distinguishing HRFI from LRFI animals. Because these panels consist of just two metabolites, it is possible to construct very fast (<5 minute/sample) and inexpensive (<$10) NMR or MS-based assays that could be used to perform bovine RFI characterization. RFI is a difficult and time-consuming measure to perform. The cost of performing RFI measurements over 80-90 days is ~ $250/head which is much higher than the cost of a metabolite test ($5-10/head) or the net profit of selecting for LRFI cattle via GrowSafe™ RFI measurements. Therefore, a simple blood test that could distinguish high RFI (HRFI) animals from LRFI animals (early on) would potentially benefit beef farmers in terms of optimizing production or selecting which animals to cull or which animals should be bred. These results suggest that serum metabolites could be used to inexpensively predict bovine RFI groupings. Further validation using a larger and diverse cohort of cattle is required to confirm these findings. After my formal study was completed, I conducted an investigation to further understand why the concentration of these metabolites in serum were significantly different between HRFI vs. LRFI bulls, especially formate which is produced by rumen microbes and used as a precursor of methane production. Hence, I measured concentration of these metabolites in ruminal fluid of the bulls but found no connection between their concentration in ruminal fluid

versus serum (data not shown). This was supported by the study of Ault-Seay et al. [373], where they found no relationship between concentration of serum and rumen metabolites in response to different dietary treatments in beef steers. Hence, I speculated that since HRFI animals produce more ruminal methane, they also produce more extra ruminal formate compared to LRFI animals. Even though concentration of formate was not significant in ruminal fluid comparing LRFI vs. HRFI bulls, the significant differentiation was later depicted in blood stream.

Hypothesis 3/Chapter 4. Approximately 70% of the cost of beef production is impacted by feed utilization. Maximizing production efficiency of beef cattle requires not only genetic selection to maximize feeding efficiency (i.e. residual feed intake - RFI), but also adequate nutrition throughout all stages of growth and development to maximize productivity of growth and reproductive capacity, even during gestation. Nutrient restriction during gestation has been shown to negatively affect post-natal growth and development. This, when combined with efforts to minimize RFI, may significantly affect energy partitioning in the offspring and subsequently important performance traits. Therefore, in chapter 4, I hypothesized that differences in selection for parental RFI and/or prenatal nutrition could change gene expression patterns in tissues of young Angus bulls leading to changes in inflammatory response and muscle development. To test this hypothesis we decided to investigate the impacts of prenatal under-nutrition during the first half of gestation (dams were either fed ~75% of NRC requirements (low-diet group (Ldiet)) vs. 100% NRC requirements (normal-diet group (Ndiet), for growing pregnant heifers during the first half of gestation), as well as selection for divergent RFI, upon gene expression in select tissues, including LT and SM muscles, liver, and testis, in Angus bull progeny. The genes monitored in this study included metabolic drivers of animal growth, especially muscle growth and development, and those that were aspects of inflammatory response. The results showed that the

mRNA abundance of protocadherin 19 [*PCDH19*] in liver, and myocyte enhancer factor 2A [*MEF2A*] in LT muscle, was significantly higher in Ldiet, and Ndiet prenatal maternal diet groups, respectively. The results also showed that in all four tissues, expression of *MEF2A* was significantly higher in the LRFI group as compared to its HRFI counterpart. We also reported correlations between gene expression in these four tissues with phenotypic measures of feed efficiency and body weight. These findings help to understand the underlying biological mechanisms regulating postnatal responses to prenatal nutrition, and feed efficiency, intake, and growth traits in beef bulls.

Hypothesis 4/Chapter 5. Prenatal exposures and intrauterine stressors in parents are known to affect the health and viability of their offspring, although the multitude of mechanisms of transfer from parent to offspring are still unclear. One of the mechanisms of transfer is thought to be through the transfer of DNA methylation patterns via the sperm. Therefore, in chapter 5, I hypothesized that differences in selection for parental RFI and/or prenatal nutrition can change DNA methylation patterns in the sperm of young Angus bulls. To test this hypothesis, WGBS was used to investigate DNA methylation patterns in the sperm of offspring Angus bulls. Dams of these bulls were either fed a Ldiet or Ndiet during the first half of gestation. We also imposed a genetic component to the experiment as half the pregnant heifers in each prenatal diet treatment were HRFI heifers bred to HRFI bulls, and the other half were LRFI heifers bred to LRFI bulls. Comparison between Ndiet and Ldiet groups identified 652 DMRs (corresponds to 352 unique DMR-associated genes) in the sperm of the bull offspring. Comparison between HRFI and LRFI parental groups identified 1400 DMRs (corresponds to 656 unique DMR-associated genes) in the sperm of the bull offspring. Through pathway analysis of the identified DMRs using the IPA tool, three networks associated with “cell survival and growth”, “disease or abnormalities”, and “connective tissue

development" were identified as being overrepresented in the DMR-associated genes when comparing the Ndiet group to the Ldiet group. Similar pathway analysis for the HRFI and LRFI bulls showed overrepresentation of the number of DMR-associated genes in four networks involved in "embryonic development", "DNA replication, DNA repair, and RNA processing", "growth control and homeostasis", as well as "lipid metabolism". These data suggest that the epigenome of sperm cells in progeny bulls are under the influence of both the prenatal environment (limited prenatal nutrition), and genetics (strong parental RFI selection), which raise questions about the inheritance of genetic potential and the potential adverse multigenerational effects of prenatal nutrition in beef cattle.

In conclusion, Ldiet bulls grew faster between 10 and 16 months of age than Ndiet bulls. From our gene expression study, we found that differential expression of genes associated with growth and development in tissues might be the reason for this growth difference between Ldiet and Ndiet bulls. This was also supported through the sperm DNA methylation study, where we found a DMR-associated gene network related to cell growth, survival, and proliferation. We also did not see any phenotypic difference in terms of fertility parameters between Ldiet and Ndiet bulls, however all three DMR-associated gene networks identified for prenatal diet contained DMR-associated genes that could be related to spermatogenesis and male fertility. Therefore, further investigation is needed to figure out whether prenatal undernutrition can threaten the fertility of bulls or not. Regarding the RFI groups, there was no growth difference between HRFI and LRFI bulls. However, we found different expression pattern of genes associated with growth and development in tissues of HRFI vs. LRFI bulls. This was also supported by the sperm DNA methylation study, where we found two DMR-associated gene networks related to growth and development, comparing HRFI vs LRFI bulls. Furthermore, through all three omics studies

(transcriptomics, epigenomics, and metabolomics) we found that LRFI and HRFI bulls were different in terms of immune response (i.e. oxidative stress). Moreover, the LRFI bulls had smaller testis, delayed onset of puberty and tended to exhibit lower progressive motility compared to HRFI bulls. This was further supported by our sperm DNA methylation study where DMR-associated gene networks containing genes related to spermatogenesis and male fertility were identified from the comparison between HRFI and LRFI bulls.

Insights from multidisciplinary omics. Each chapter on its own gives insight into the biological processes affected by prenatal maternal nutrition, and selection for and/or expression of RFI. However, greater knowledge is gained when investigating the results in the light of the other studies. To this end, a relationship between metabolome and transcriptome concerning RFI was further investigated. In Figure 6.1, I have shown the relationship between *MEF2A* gene and significant serum metabolites of RFI in a network, in which you see a connection to Ca2+, proinsulin, three amino acids, formic acid, LysoPC(28:0) and three acylcarnitines. The *MEF2* family of genes responds to multiple calcium-regulated signals, in the control of skeletal muscle fiber type [374]. These Ca2+ channels are also responsible for regulating secretion of insulin in Langerhans islets [375]. Insulin stimulates protein synthesis in muscle tissue when the levels of total amino acids, or at least the essential amino acids, are at or above their post-absorptive concentrations [376]. As well, a study conducted by Nascimento et al. [377] reported higher concentration of insulin in the plasma of LRFI Nellore cattle, which in retrospect, would have been interesting to investigate in our bulls. As alluded to by figure 6.1, BCAAs (leucine, iso-leucine, and valine) are essential amino acids having the primary role in stimulating muscle protein synthesis [376]. Serine and glycine are two amino acids that are biosynthetically linked, and together provide the essential precursors for the synthesis of proteins [378-380]. Therefore, there

may be a concrete biological connection between *MEF2A* expression and concentrations of leucine, serine and glycine in the serum of LRFI bulls.

It has also been shown that high concentrations of BCAAs are associated with higher oxidative stress, and as seen in human studies, can serve as biomarkers for obesity-associated insulin resistance and diabetes [381]. Acylcarnitines mediate activation of several important hepatic metabolic signaling pathways leading to diseases such as non-alcoholic fatty liver disease and type 2 diabetes mellitus [230, 231]. Formate is the simplest carboxylic acid and potent reductive force against oxidative stress [242]. In an oxidative milieu, glycine is catabolized to form glyoxylate in order to combat ROS that appear in cells [242]. Decreased oxidative stress in the liver is associated with lower feed maintenance requirements, due to a lower lipid and protein turnover and better efficiency in energy usage [241]. A recent study conducted by Casal et al. reported lower levels of ROS in the liver of LRFI steers which suggests they have lower levels of hepatic oxidative stress than HRFI steers [241]. Therefore, lower level of glycine and higher level of formate in the serum of HRFI animals suggest that these less efficient cattle are more prone to oxidative stress. Therefore in our study, expression of *MEF2A* was higher in LRFI group, and could be reacting to the different levels of oxidative stress in high versus low RFI animals, at least in liver tissue.

Future Directions. The application of metabolomics for identifying and quantifying metabolite profile of different biofluids and tissues is relatively new and challenges still exist with both the technology and the experimental protocols. As seen in chapter 2, we could confidently infer or predict the existence of 51,801 unique metabolites in the bovine metabolome, our experimentally acquired data corresponded to just 7.4% of the total number of metabolites/chemicals reported in the BMDB. Clearly more work is needed in the field to expand

the proportion of the metabolome that can be routinely measured in metabolomics studies. Because of their fundamentally different separation and detection technologies, different metabolomics platforms tend to target or detect different classes of metabolites. For instance, NMR is relatively "untargeted" but is biased towards highly abundant, water-soluble compounds. Other methods were quite targeted, with ICP-MS being limited only to metal ions and LC-MS/MS (the TMIC Prime assay) being limited to a pre-selected set of 143 compounds including amino acids, biogenic amines, organic acids, lipid-like compounds. While the metabolomic studies described in chapters 2 and 3 did employ a relatively wide range of metabolomics platforms (NMR, LC-MS/MS, ICP-MS), they did not use all available detection tools (GC-MS, GC×GC-TOF) nor did they explore all available separation protocols (e.g., solid phase extraction and enrichment, chemical derivatization, etc.). Therefore, in the future, the application of other metabolomics techniques, such as GC-MS, GC×GC-MS (to help identify volatile compounds), untargeted high resolution-mass spectrometry (HR-MS), and more extensive targeted LC-MS/MS techniques could and should be used to extend the metabolite coverage of bovine biofluids and tissues.

As seen in chapters 3, 4 and 5, the main limitation of each omics study was the relatively small sample size. Omics studies typically need sample sized of dozens to hundreds of animals or samples to draw strong conclusions. While the findings made in the smaller-scale pilot studies described in this thesis are statistically significant, it is clear that these studies will need to be repeated on larger animal cohorts to confirm their validity. Likewise, these expanded studies would have to include more diverse animal cohorts covering a broader range with regard to diet, sex, age, breed, and farm management practices. Addressing these limitations in future studies should help confirm the results reported here and improve our understanding of the underlying biological mechanisms regulating postnatal responses to prenatal nutrition, and feed efficiency, intake, and

growth traits in beef cattle. Addressing these limitations would also allow greater integration and more expanded interpretation of the different omics studies that were attempted in this thesis.

Overall with regards to chapter 2, the results presented have implications far beyond the field of metabolomics, especially given the economic importance of the bovine metabolome in the food industry and its importance in human nutrition. We expect these data to serve as a benchmark in comparing various technologies and assessing future methodological improvements in bovine metabolome research. In the meantime, it is hoped that the BMDB will provide a reliable source for metabolomics researchers, animal scientists, food chemists, nutritional scientists, and consumers by providing a comprehensive, easy-to-use and highly centralized web-based resource on the chemical composition of bovine biofluids and tissues. For example in the beef industry, by measuring the blood metabolites (i.e. minerals) from cattle and comparing them with the normal ranges of metabolites available in BMDB, we can understand the health condition of an animal. Regarding chapter 3, we hope that once the two biomarker panels were verified in a larger and more diverse animal cohort, beef producers can use these biomarkers to inexpensively classify their cattle into LRFI and HRFI groups. Chapters 4 and 5 also help to better understand the biology behind the impacts of prenatal undernutrition and RFI on young Angus bulls. These two chapters provided some evidence that prenatal undernutrition negatively affects growth and development as well as health of offspring cattle. More specifically in chapter 4, we found that expression of *PCDH19* (associated with immune function) and *MEF2A* (associated with muscle development) genes were significantly affected by prenatal undernutrition which somehow explains how prenatal undernutrition might negatively affect the overall health and growth of offspring beef bulls. As well, higher expression of *MEF2A* in four tissues of LRFI bulls might answer the mechanism of why LRFI animals are leaner. Chapter 5 also highlights how both prenatal undernutrition and

parental RFI can affect epigenetic DNA machinery in the sperm of next generation which can be passed to the next generation. Overall, this research provided valuable outcomes that can be used by animal scientists, food chemists, nutritional scientists and other researchers.

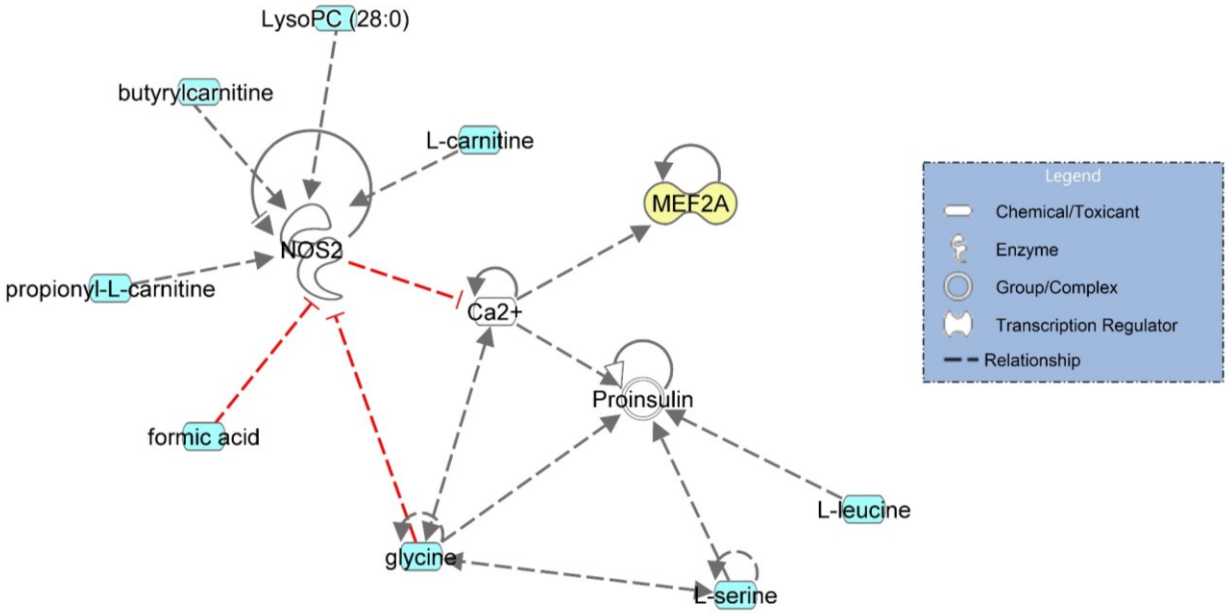

Figure 6. 1 An example of the relationship between metabolome and transcriptome with regards to RFI found in this study

REFERENCES

1. Canfax Research Services Production & Management Practices that Drive Profitability. **2017**.

2. Kaliel, D. A. Insights into Managing Winter Feed Costs in Alberta Cow/calf Operations. **2004**.

3. Du, M.; Tong, J.; Zhao, J.; Underwood, K. R.; Zhu, M.; Ford, S. P.; Nathanielsz, P. W. Fetal programming of skeletal muscle development in ruminant animals. *J. Anim. Sci.* **2010***, 88*, 51.

4. Funston, R. N.; Larson, D. M.; Vonnahme, K. A. Effects of maternal nutrition on conceptus growth and offspring performance: implications for beef cattle production. *J. Anim. Sci.* **2010***, 88*, 205.

5. Long, N. M.; Vonnahme, K. A.; Hess, B. W.; Nathanielsz, P. W.; Ford, S. P. Effects of early gestational undernutrition on fetal growth, organ development, and placentomal composition in the bovine. *J. Anim. Sci.* **2009***, 87*, 1950-1959.

6. Long, N. M.; Tousley, C. B.; Underwood, K. R.; Paisley, S. I.; Means, W. J.; Hess, B. W.; Du, M.; Ford, S. P. Effects of early- to mid-gestational undernutrition with or without protein supplementation on offspring growth, carcass characteristics, and adipocyte size in beef cattle. *J. Anim. Sci.* **2012***, 90*, 197-206.

7. Basarab, J. A.; Beauchemin, K. A.; Baron, V. S.; Ominski, K. H.; Guan, L. L.; Miller, S. P.; Crowley, J. J. Reducing GHG emissions through genetic improvement for feed efficiency: effects on economically important traits and enteric methane production. *Animal* **2013***, 7*, 303-315.

8. Crews, D. H. Genetics of efficient feed utilization and national cattle evaluation: a review. *Genet. Mol. Res.* **2005***, 4*, 152-165.

9. Carstens GE, Tedeschi LO. 2006. Defining feed efficiency in beef cattle. Proc Beef Improv Federation 38th Ann Res Symp Ann Meet Choctow, MS, USA, 12-21.

10. Bishop, M. D.; Davis, M. E.; Harvey, W. R.; Wilson, G. R.; VanStavern, B. D. Divergent selection for postweaning feed conversion in Angus beef cattle: II. Genetic and phenotypic correlations and realized heritability estimate. *J. Anim. Sci.* **1991***, 69*, 4360-4367.

11. Archer, J. A.; Richardson, E. C.; Herd, R. M.; Arthur, P. F. Potential for selection to improve efficiency of feed use in beef cattle: a review. *Aust. J. Agric. Res.* **1999***, 50*, 147-162.

12. Herd, R. M.; Bishop, S. C. Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. *Livest. Prod. Sci.* **2000***, 63*, 111-119.

13. Kelly, A. K.; McGee, M.; Crews, D. H.; Fahey, A. G.; Wylie, A. R.; Kenny, D. A. Effect of divergence in residual feed intake on feeding behavior, blood metabolic variables, and body composition traits in growing beef heifers. *J. Anim. Sci.* **2010***, 88*, 109-123.

14. Koch, R. M.; Swiger, L. A.; Chambers, D.; Gregory, K. E. Efficiency of feed use in beef cattle. *J. Anim. Sci.* **1963***, 22*, 486-494.

15. Arthur, P. F.; Herd, R. M.; Wright, J.; Xu, G.; Dibley, K.; Richardson, E. C. Net feed conversion efficiency and its relationship with other traits in beef cattle. *Proc Aust Soc Anim Prod* **1996***, 21*, 107-110.

16. Basarab, J. A.; Price, M. A.; Aalhus, J. L.; Okine, E. K.; Snelling, W. M.; Lyle, K. L. Residual feed intake and body composition in young growing cattle. *Can. J. Anim. Sci.* **2003***, 83*, 189-204.

17. Nkrumah, J. D.; Okine, E. K.; Mathison, G. W.; Schmid, K.; Li, C.; Basarab, J. A.; Price, M. A.; Wang, Z.; Moore, S. S. Relationships of feedlot feed efficiency, performance, and

feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J. Anim. Sci.* **2006***, 84*, 145-153.

18. Byerly TC. 1941. Feed and other costs of producing market eggs. Tech. Bull. n.1. The University of Maryland Agric. Exp., College Park, Maryland.

19. Herd, R. M.; Arthur, P. F. Physiological basis for residual feed intake. *J. Anim. Sci.* **2009***, 87*, 64.

20. Herd, R. M.; Oddy, V. H.; Richardson, E. C. Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms. *Aust. J. Exp. Agric.* **2004***, 44*, 423-430.

21. Kelly, A. K.; Waters, S. M.; McGee, M.; Fonseca, R. G.; Carberry, C.; Kenny, D. A. mRNA expression of genes regulating oxidative phosphorylation in the muscle of beef cattle divergently ranked on residual feed intake. *Physiol. Genomics* **2011***, 43*, 12-23.

22. Arthur, P. F.; Archer, J. A.; Johnston, D. J.; Herd, R. M.; Richardson, E. C.; Parnell, P. F. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *J. Anim. Sci.* **2001***, 79*, 2805-2811.

23. Fitzsimons, C.; Kenny, D. A.; Deighton, M. H.; Fahey, A. G.; McGee, M. Methane emissions, body composition, and rumen fermentation traits of beef heifers differing in residual feed intake. *J. Anim. Sci.* **2013***, 91*, 5789-5800.

24. Herd, R. M.; Hegarty, R. S.; Dicker, R. W.; Archer, J. A.; Arthur, P. F. Selection for residual feed intake improves feed conversion in steers on pasture. *Proc. Aust. Soc. Anim. Prod.* **2002***, 24*, 85-88.

25. Beauchemin, K. A.; Henry Janzen, H.; Little, S. M.; McAllister, T. A.; McGinn, S. M. Life cycle assessment of greenhouse gas emissions from beef production in western Canada: A case study. *Agricultural Systems* **2010***, 103*, 371-379.

26. Kennedy, B. W.; van der Werf, J H; Meuwissen, T. H. Genetic and statistical properties of residual feed intake. *J. Anim. Sci.* **1993***, 71*, 3239-3250.

27. Herd, R. M.; Arthur, P. F.; Donoghue, K. A.; Bird, S. H.; Bird-Gardiner, T.; Hegarty, R. S. Measures of methane production and their phenotypic relationships with dry matter intake, growth, and body composition traits in beef cattle. *J. Anim. Sci.* **2014***, 92*, 5267-5274.

28. Nkrumah, J. D.; Crews, D. H.; Basarab, J. A.; Price, M. A.; Okine, E. K.; Wang, Z.; Li, C.; Moore, S. S. Genetic and phenotypic relationships of feeding behavior and temperament with performance, feed efficiency, ultrasound, and carcass merit of beef cattle. *J. Anim. Sci.* **2007***, 85*, 2382-2390.

29. Castro Bulle, F. C.; Paulino, P. V.; Sanches, A. C.; Sainz, R. D. Growth, carcass quality, and protein and energy metabolism in beef cattle with different growth potentials and residual feed intakes. *J. Anim. Sci.* **2007***, 85*, 928-936.

30. Richardson, E. C.; Herd, R. M.; Oddy, V. H.; Thompson, J. M.; Archer, J. A.; Arthur, P. F. Body composition and implications for heat production of Angus steer progeny of parents selected for and against residual feed intake. *Aust. J. Exp. Agric.* **2001***, 41*, 1065-1072.

31. Bishop, S. C. Phenotypic and genetic variation in body weight, food intake and energy utilisation in Hereford cattle II. Effects of age and length of performance test. *Livest. Prod. Sci.* **1992***, 30*, 19-31.

32. Carstens GE, Johnson DE, Johnson KA, Hotovy SK, Szymanski TJ. 1989. Genetic variation in energy expenditures of monozygous twin beef cattle at 9 and 20 months of age. In Energy

metabolism of farm animals: Proc 11th Symp. Lutenen, The Netherlands. EAPP Publication No. 43.

33. Wang, Z.; Nkrumah, J. D.; Li, C.; Basarab, J. A.; Goonewardene, L. A.; Okine, E. K.; Crews, D. H.; Moore, S. S. Test duration for growth, feed intake, and feed efficiency in beef cattle using the GrowSafe System. *J. Anim. Sci.* **2006***, 84*, 2289-2298.

34. Standing Committee on Agriculture. 2000. Feeding standards for Australian livestock. Ruminants. CSIRO Publications, East Melbourne, Australia.

35. Basarab, J. A.; McCartney, D.; Okine, E. K.; Baron, V. S. Relationships between progeny residual feed intake and dam productivity traits. *Can. J. Anim. Sci.* **2007***, 87*, 489-502.

36. Durunna, O. N.; Mujibi, F. D.; Goonewardene, L.; Okine, E. K.; Basarab, J. A.; Wang, Z.; Moore, S. S. Feed efficiency differences and reranking in beef steers fed grower and finisher diets. *J. Anim. Sci.* **2011***, 89*, 158-167.

37. Basarab, J. A.; Colazo, M. G.; Ambrose, D. J.; Novak, S.; McCartney, D.; Baron, V. S. Residual feed intake adjusted for backfat thickness and feeding frequency is independent of fertility in beef heifers. *Can. J. Anim. Sci.* **2011***, 91*, 573-584.

38. Manafiazar, G.; Basarab, J. A.; Baron, V. S.; McKeown, L.; Doce, R. R.; Swift, M.; Undi, M.; Wittenberg, K.; Ominski, K. Effect of post-weaning residual feed intake classification on grazed grass intake and performance in pregnant beef heifers. *Can. J. Anim. Sci.* **2015***, 95*, 369-381.

39. McDonald, T. J.; Nichols, B. M.; Harbac, M. M.; Norvell, T. M.; Paterson, J. Dry matter intake is repeatable over parities and residual feed intake is negatively correlated with dry matter digestibility in gestating cows. *J. Anim. Sci.* **2010***, 88*.

40. Krueger, W. K.; Carstens, G.; Lancaster, P.; Slay, L. J.; Miller, J. C.; Forbes, T. D. A. Relationship between residual feed intake and apparent nutrient digestibility in growing calves. *J. Anim. Sci.* **2008***, 86*.

41. Gomes, R. d. C.; Sainz, R. D.; Leme, P. R. Protein metabolism, feed energy partitioning, behavior patterns and plasma cortisol in Nellore steers with high and low residual feed intake. *Revista Brasileira de Zootecnia* **2013***, 42*, 44-50.

42. Lawrence, P.; Kenny, D. A.; Earley, B.; Crews, D. H.; McGee, M. Grass silage intake, rumen and blood variables, ultrasonic and body measurements, feeding behavior, and activity in pregnant beef heifers differing in phenotypic residual feed intake. *J. Anim. Sci.* **2011***, 89*, 3248-3261.

43. Carstens, G.; Theis, C. M.; White, M. B.; Jr, T. H.; Warrington, B. G.; Randel, R.; Forbes, T. D. A.; Lippke, H.; Greene, L.; Lunt, D. Residual feed intake in beef steers: I. Correlations with performance traits and ultrasound measures of body composition. *Proc. West. Sect. Am. Soc. Anim. Sci.* **2002***, 53*, 552-555.

44. Schenkel, F. S.; Miller, S. P.; Wilton, J. W. Genetic parameters and breed differences for feed efficiency, growth, and body composition traits of young beef bulls. *Can. J. Anim. Sci.* **2004***, 84*, 177-185.

45. Arthur, P. F.; Archer, J. A.; Herd, R. M.; Richardson, E. C.; Wright, J. H.; Dibley, K. C. P.; Burton, D. A. Genotypic and phenotypic variation in feed intake, feed efficiency and growth rate in beef cattle. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* **1997***, 12*, 234-237.

46. Cruzen, S. M.; Harris, A. J.; Hollinger, K.; Punt, R. M.; Grubbs, J. K.; Selsby, J. T.; Dekkers, J. C.; Gabler, N. K.; Lonergan, S. M.; Huff-Lonergan, E. Evidence of decreased muscle

protein turnover in gilts selected for low residual feed intake. *J. Anim. Sci.* **2013***, 91*, 4007-4016.

47. Lefaucheur, L.; Lebret, B.; Ecolan, P.; Louveau, I.; Damon, M.; Prunier, A.; Billon, Y.; Sellier, P.; Gilbert, H. Muscle characteristics and meat quality traits are affected by divergent selection on residual feed intake in pigs. *J. Anim. Sci.* **2011***, 89*, 996-1010.

48. de Haer, L. C. M.; Luiting, P.; Aarts, H. L. M. Relations among individual (residual) feed intake, growth performance and feed intake pattern of growing pigs in group housing. *Livest. Prod. Sci.* **1993***, 36*, 233-253.

49. Luiting, P.; Schrama, J. W.; van der Hel, W.; Urff, E. M. Metabolic differences between White Leghorns selected for high and low residual food consumption. *Br. Poult. Sci.* **1991***, 32*, 763-782.

50. Mousel, M. R.; Stroup, W. W.; Nielsen, M. K. Locomotor activity, core body temperature, and circadian rhythms in mice selected for high or low heat loss. *J. Anim. Sci.* **2001***, 79*, 861-868.

51. Bunger, L.; Macleod, M. G.; Wallace, C. A.; Hill, W. G. Direct and correlated effects of selection for food intake corrected for body weight in the adult mouse. *Proceedings of the World Congress on Genetics Applied to Livestock Production* **1998***, 26*, 97–100.

52. Richardson EC, Herd RM, Oddy VH. 2000. Variation in body composition, activity and other physiological processes and their associations with feed efficiency. Feed efficiency in beef cattle. In: JA Archer JA, Herd RM, Arthur PF (Eds.) Proceedings of the feed efficiency workshop. University of New England: Armidale, NSW. pp.46-50.

53. Britt, J. S.; Thomas, R. C.; Speer, N. C.; Hall, M. B. Efficiency of converting nutrient dry matter to milk in Holstein herds. *J. Dairy Sci.* **2003***, 86*, 3796-3801.

54. DiGiacomo, K.; Marett, L. C.; Wales, W. J.; Hayes, B. J.; Dunshea, F. R.; Leury, B. J. Thermoregulatory differences in lactating dairy cattle classed as efficient or inefficient based on residual feed intake. *Anim. Prod. Sci.* **2014***, 54*, 1877-1881.

55. Barea, R.; Dubois, S.; Gilbert, H.; Sellier, P.; van Milgen, J.; Noblet, J. Energy utilization in pigs selected for high and low residual feed intake. *J. Anim. Sci.* **2010***, 88*, 2062-2072.

56. Martello, L. S.; da Luz E Silva, S; da Costa Gomes, R.; da Silva Corte, R R; Leme, P. R. Infrared thermography as a tool to evaluate body surface temperature and its relationship with feed efficiency in Bos indicus cattle in tropical conditions. *Int. J. Biometeorol.* **2016***, 60*, 173-181.

57. Intergovernmental Panel on Climate Change (IPCC). 2014. Global Warming Potential Values. URL: https://www.ipcc.ch/report/ar5/.

58. Steinfeld, H.; Wassenaar, T. The Role of Livestock Production in Carbon and Nitrogen Cycles. *Annu. Rev. Environ. Resour.* **2007***, 32*, 271-294.

59. Intergovernmental Panel on Climate Change (IPCC). 2006. Revised guidelines for national greenhouse gas inventories. IPCC/OECD/IEA/IGES. Vol. 4. Agriculture, forestry and other land use. Chapter 11. N2O emissions from managed soils and CO2 emissions from lime and urea application. URL: http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/4-Volume_4/V4._11-Ch11_N2O&CO2pdfhtm.

60. Russell RW, Gahr SA 2000. Glucose availability and associated metabolism. In: D'Mello JPF (Ed.) Farm animal metabolism and nutrition (ed.) Centre for Agricultural Bioscience International Publishing, Wallingford, UK. pp.121-147.

61. Hernandez-Sanabria, E.; Guan, L. L.; Goonewardene, L. A.; Li, M.; Fujibi, D.; Stothard, P.; Moore, S. S.; Leon-Quintero, M. C. Association between microbial diversity and microbial

fermentation parameters in the bovine rumen and host's feed efficiency traits. *Appl Environ Microbiol* **2010***, 76*, 6338-6350.

62. Guan, L. L.; Nkrumah, J. D.; Basarab, J. A.; Moore, S. S. Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. *FEMS Microbiol. Lett.* **2008***, 288*, 85-91.

63. Carberry, C. A.; Kenny, D. A.; Han, S.; McCabe, M. S.; Waters, S. M. Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. *Appl. Environ. Microbiol.* **2012***, 78*, 4949-4958.

64. Grubbs, J. K.; Fritchen, A. N.; Huff-Lonergan, E.; Gabler, N. K.; Lonergan, S. M. Selection for residual feed intake alters the mitochondria protein profile in pigs. *J. Proteomics* **2013***, 80*, 334-345.

65. Colpoys JD. 2015. Swine feed efficiency: Implications for swine behavior, physiology and welfare. PhD Dissertation, Iowa State University.

66. Rakhshandeh, A.; Dekkers, J. C.; Kerr, B. J.; Weber, T. E.; English, J.; Gabler, N. K. Effect of immune system stimulation and divergent selection for residual feed intake on digestive capacity of the small intestine in growing pigs. *J. Anim. Sci.* **2012***, 90 Suppl 4*, 233-235.

67. Azarpajouh, S.; Colpoys, J.; Rakhshandeh, A.; Dekkers, J.; Bruns, C.; Gabler, N.; Johnson, A. Residual feed intake selection: Effect on gilt behavior in response to a lipopolysaccharide challenge. **2015**.

68. Dunkelberger, J. R.; Boddicker, N. J.; Serão, N. V. L.; Young, J. M.; Rowland, R. R. R.; Dekkers, J. C. M. Response of pigs divergently selected for residual feed intake to experimental infection with the PRRS virus. *Livestock Science* **2015***, 177*, 132-141.

69. Consolo, N. R. B.; Munro, J. C.; Bourgon, S. L.; Karrow, N. A.; Fredeen, A. H.; Martell, J. E.; Montanholi, Y. R. Associations of blood analysis with feed efficiency and developmental stage in grass-fed beef heifers. *Animals (Basel)* **2018**, *8*, 10.3390/ani8080133.

70. Herd, R. M.; Velazco, J. I.; Smith, H.; Arthur, P. F.; Hine, B.; Oddy, H.; Dobos, R. C.; Hegarty, R. S. Genetic variation in residual feed intake is associated with body composition, behavior, rumen, heat production, hematology, and immune competence traits in Angus cattle1. *J. Anim. Sci.* **2019**, *97*, 2202-2219.

71. Dickerson, G. Efficiency of Animal Production—Molding the Biological Components. *J. Anim. Sci.* **1970**, *30*, 849-859.

72. Smith BA. 2012. The assessment of replacement heifer production efficiencies through residual feed intake and key hormone profiles. Master's thesis, The University of Guelph. URL: https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/5212/Smith_Brock_201301_Msc.pdf?sequence=3.

73. Johnson, C.; Fitzsimmons, C.; Colazo, M.; Li, C.; Kastelic, J.; Thundathil, J. Impacts of residual feed intake and pre-natal diet on reproductive potential of bulls. *Anim. Prod. Sci.* **2019**, *59*, 1827-1836.

74. Randel, R. D.; Welsh, T. H. Joint Alpharma-Beef Species Symposium: interactions of feed efficiency with beef heifer reproductive development. *J. Anim. Sci.* **2013**, *91*, 1323-1328.

75. Shaffer, K. S.; Turk, P.; Wagner, W. R.; Felton, E. E. Residual feed intake, body composition, and fertility in yearling beef heifers. *J. Anim. Sci.* **2011**, *89*, 1028-1034.

76. Wang, Z.; Colazo, M. G.; Basarab, J. A.; Goonewardene, L. A.; Ambrose, D. J.; Marques, E.; Plastow, G.; Miller, S. P.; Moore, S. S. Impact of selection for residual feed intake on breeding soundness and reproductive performance of bulls on pasture-based multisire mating. *J. Anim. Sci.* **2012***, 90*, 2963-2969.

77. Awda, B. J.; Miller, S. P.; Montanholi, Y. R.; Voort, G. V.; Caldwell, T.; Buhr, M. M.; Swanson, K. C. The relationship between feed efficiency traits and fertility in young beef bulls. *Can. J. Anim. Sci.* **2013***, 93*, 185-192.

78. Montanholi, Y. R.; Fontoura, A. B.; Diel de Amorim, M.; Foster, R. A.; Chenier, T.; Miller, S. P. Seminal plasma protein concentrations vary with feed efficiency and fertility-related measures in young beef bulls. *Reprod. Biol.* **2016***, 16*, 147-156.

79. Fontoura, A. B.; Montanholi, Y. R.; Diel de Amorim, M.; Foster, R. A.; Chenier, T.; Miller, S. P. Associations between feed efficiency, sexual maturity and fertility-related measures in young beef bulls. *Animal* **2016***, 10*, 96-105.

80. de Rooij, S. R.; Painter, R. C.; Roseboom, T. J.; Phillips, D. I.; Osmond, C.; Barker, D. J.; Tanck, M. W.; Michels, R. P.; Bossuyt, P. M.; Bleker, O. P. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. *Diabetologia* **2006***, 49*, 637-643.

81. Roseboom, T. J.; Painter, R. C.; van Abeelen, A. F.; Veenendaal, M. V.; de Rooij, S. R. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas* **2011***, 70*, 141-145.

82. Barker, D. J.; Eriksson, J. G.; Forsen, T.; Osmond, C. Fetal origins of adult disease: strength of effects and biological basis. *Int. J. Epidemiol.* **2002***, 31*, 1235-1239.

83. Caton, J.; Hess, B. Maternal plane of nutrition: Impacts on fetal outcomes and postnatal offspring responses. *Proceedings of 4th Grazing Livestock Nutrition Conference* **2010**, 104-122.

84. Reynolds, L. P.; Caton, J. S. Role of the pre- and post-natal environment in developmental programming of health and productivity. *Mol. Cell. Endocrinol.* **2012***, 354*, 54-59.

85. Du, M.; Zhao, J. X.; Yan, X.; Huang, Y.; Nicodemus, L. V.; Yue, W.; McCormick, R. J.; Zhu, M. J. Fetal muscle development, mesenchymal multipotent cell differentiation, and associated signaling pathways. *J. Anim. Sci.* **2011***, 89*, 583-590.

86. Reynolds, L. P.; Vonnahme, K. A.; Lemley, C. O.; Redmer, D. A.; Grazul-Bilska, A. T.; Borowicz, P. P.; Caton, J. S. Maternal stress and placental vascular function and remodeling. *Curr. Vasc. Pharmacol.* **2013***, 11*, 564-593.

87. Vonnahme, K. A.; Lemley, C. O.; Caton, J. S.; Meyer, A. M. Impacts of Maternal Nutrition on Vascularity of Nutrient Transferring Tissues during Gestation and Lactation. *Nutrients* **2015***, 7*, 3497-3523.

88. Zhang, S.; Regnault, T. R.; Barker, P. L.; Botting, K. J.; McMillen, I. C.; McMillan, C. M.; Roberts, C. T.; Morrison, J. L. Placental adaptations in growth restriction. *Nutrients* **2015***, 7*, 360-389.

89. Swanson, T. J.; Hammer, C. J.; Luther, J. S.; Carlson, D. B.; Taylor, J. B.; Redmer, D. A.; Neville, T. L.; Reed, J. J.; Reynolds, L. P.; Caton, J. S.; Vonnahme, K. A. Effects of gestational plane of nutrition and selenium supplementation on mammary development and colostrum quality in pregnant ewe lambs. *J. Anim. Sci.* **2008***, 86*, 2415-2423.

90. Reed, J. J.; Ward, M. A.; Vonnahme, K. A.; Neville, T. L.; Julius, S. L.; Borowicz, P. P.; Taylor, J. B.; Redmer, D. A.; Grazul-Bilska, A. T.; Reynolds, L. P.; Caton, J. S. Effects of

selenium supply and dietary restriction on maternal and fetal body weight, visceral organ mass and cellularity estimates, and jejunal vascularity in pregnant ewe lambs. *J. Anim. Sci.* **2007***, 85*, 2721-2733.

91. NRC. 2007. Nutrients requirements of small ruminants, sheep, goats, cervids and new world camelids. Washington (DC): National Academic Press.

92. NRC. 1996. Nutrient requirements of beef cattle, 7th rev. ed. Washington (DC): National Academic Press.

93. Robinson, J. J.; Sinclair, K. D.; McEvoy, T. G. Nutritional effects on foetal growth. *Animal Science* **1999***, 68*, 315-331.

94. NASEM. 2016. Nutrient requirements of beef cattle, 8th rev. ed. Washington (DC): National Academic Press.

95. Greenwood, P. L.; Hunt, A. S.; Hermanson, J. W.; Bell, A. W. Effects of birth weight and postnatal nutrition on neonatal sheep: II. Skeletal muscle growth and development. *J. Anim. Sci.* **2000***, 78*, 50-61.

96. Zhu, M. J.; Ford, S. P.; Means, W. J.; Hess, B. W.; Nathanielsz, P. W.; Du, M. Maternal nutrient restriction affects properties of skeletal muscle in offspring. *J. Physiol.* **2006***, 575*, 241-250.

97. Ford, S. P.; Hess, B. W.; Schwope, M. M.; Nijland, M. J.; Gilbert, J. S.; Vonnahme, K. A.; Means, W. J.; Han, H.; Nathanielsz, P. W. Maternal undernutrition during early to mid-gestation in the ewe results in altered growth, adiposity, and glucose tolerance in male offspring. *J. Anim. Sci.* **2007***, 85*, 1285-1294.

98. Taylor, N. A.; Wilkinson, J. G. Exercise-induced skeletal muscle growth. Hypertrophy or hyperplasia? *Sports Med.* **1986***, 3*, 190-200.

99. McCoard, S. A.; McNabb, W. C.; Peterson, S. W.; McCutcheon, S. N.; Harris, P. M. Muscle growth, cell number, type and morphometry in single and twin fetal lambs during mid to late gestation. *Reprod. Fertil. Dev.* **2000***, 12*, 319-327.

100. Cafe, L.; Hennessy, D. W.; Hearnshaw, H.; Morris, S.; Greenwood, P. Influences of nutrition during pregnancy and lactation on birth weights and growth to weaning of calves sired by Piedmontese or Wagyu bulls. *Aust. J. Exp. Agric.* **2006***, 46*, 245-255.

101. Greenwood, P.; Cafe, L.; Hearnshaw, H.; Hennessy, D.; Morris, S. Consequences of prenatal and preweaning growth for yield of beef primal cuts from 30-month-old Piedmontese- and Wagyu-sired cattle. *Animal Production Science* **2009***, 49*, 468-478.

102. Blair, A. D.; Mohrhauser, D. A.; Taylor, A. R.; Underwood, K. R.; Pritchard, R. H.; and Wertz-Lutz, A. E. In *Preganant cow nutrition: effect on progeny carcass and meat Characteristic;* 2013, pp 318.

103. Gardner, D. S.; Tingey, K.; Van Bon, B. W.; Ozanne, S. E.; Wilson, V.; Dandrea, J.; Keisler, D. H.; Stephenson, T.; Symonds, M. E. Programming of glucose-insulin metabolism in adult sheep after maternal undernutrition. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2005***, 289*, 947.

104. Zhu, M. J.; Du, M.; Hess, B. W.; Means, W. J.; Nathanielsz, P. W.; Ford, S. P. Maternal nutrient restriction upregulates growth signaling pathways in the cotyledonary artery of cow placentomes. *Placenta* **2007***, 28*, 361-368.

105. Corah, L. R.; Dunn, T. G.; Kaltenbach, C. C. Influence of prepartum nutrition on the reproductive performance of beef females and the performance of their progeny. *J. Anim. Sci.* **1975***, 41*, 819-824.

106. Budge, H.; Bispham, J.; Dandrea, J.; Evans, E.; Heasman, L.; Ingleton, P. M.; Sullivan, C.; Wilson, V.; Stephenson, T.; Symonds, M. E. Effect of maternal nutrition on brown adipose tissue and its prolactin receptor status in the fetal lamb. *Pediatr. Res.* **2000***, 47*, 781-786.

107. Qi, Z.; Zhao, H.; Zhang, Q.; Bi, Y.; Ren, L.; Zhang, X.; Yang, H.; Yang, X.; Wang, Q.; Li, C.; Zhou, J.; Xin, Y.; Yang, Y.; Yang, H.; Du, Z.; Tan, Y.; Han, Y.; Song, Y.; Zhou, L.; Zhang, P.; Cui, Y.; Yan, Y.; Zhou, D.; Yang, R.; Wang, X. Acquisition of maternal antibodies both from the placenta and by lactation protects mouse offspring from Yersinia pestis challenge. *Clin. Vaccine Immunol.* **2012***, 19*, 1746-1750.

108. Hough, R. L.; McCarthy, F. D.; Kent, H. D.; Eversole, D. E.; Wahlberg, M. L. Influence of nutritional restriction during late gestation on production measures and passive immunity in beef cattle. *J. Anim. Sci.* **1990***, 68*, 2622-2627.

109. Rae, M. T.; Kyle, C. E.; Miller, D. W.; Hammond, A. J.; Brooks, A. N.; Rhind, S. M. The effects of undernutrition, in utero, on reproductive function in adult male and female sheep. *Anim. Reprod. Sci.* **2002***, 72*, 63-71.

110. Mossa, F.; Carter, F.; Walsh, S. W.; Kenny, D. A.; Smith, G. W.; Ireland, J. L.; Hildebrandt, T. B.; Lonergan, P.; Ireland, J. J.; Evans, A. C. Maternal undernutrition in cows impairs ovarian and cardiovascular systems in their offspring. *Biol. Reprod.* **2013***, 88*, 92.

111. Rae, M. T.; Rhind, S. M.; Fowler, P. A.; Miller, D. W.; Kyle, C. E.; Brooks, A. N. Effect of maternal undernutrition on fetal testicular steroidogenesis during the CNS androgen-responsive period in male sheep fetuses. *Reproduction* **2002***, 124*, 33-39.

112. Rinaldi, J. C.; Justulin, L. A.; Lacorte, L. M.; Sarobo, C.; Boer, P. A.; Scarano, W. R.; Felisbino, S. L. Implications of intrauterine protein malnutrition on prostate growth, maturation and aging. *Life Sci.* **2013***, 92*, 763-774.

113. Da Silva, P.; Aitken, R. P.; Rhind, S. M.; Racey, P. A.; Wallace, J. M. Influence of placentally mediated fetal growth restriction on the onset of puberty in male and female lambs. *Reproduction* **2001***, 122*, 375-383.

114. Fiehn, O. Metabolomics--the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002***, 48*, 155-171.

115. Goldansaz, S. A.; Guo, A. C.; Sajed, T.; Steele, M. A.; Plastow, G. S.; Wishart, D. S. Livestock metabolomics and the livestock metabolome: a systematic review. *PLoS One* **2017***, 12*, e0177675.

116. Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; Dame, Z. T.; Poelzer, J.; Huynh, J.; Yallou, F. S.; Psychogios, N.; Dong, E.; Bogumil, R.; Roehring, C.; Wishart, D. S. The human urine metabolome. *PLoS One* **2013***, 8*, e73076.

117. Monteiro, M. S.; Carvalho, M.; Bastos, M. L.; Guedes de Pinho, P. Metabolomics analysis for biomarker discovery: advances and challenges. *Curr. Med. Chem.* **2013***, 20*, 257-271.

118. Wilson, D. M.; Burlingame, A. L. Deuterium and carbon-13 tracer studies of ethanol metabolism in the rat by 2H, 1H-decoupled 13C nuclear magnetic resonance. *Biochem. Biophys. Res. Commun.* **1974***, 56*, 828-835.

119. Nicholson, J. K.; Lindon, J. C.; Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999***, 29*, 1181-1189.

120. Reo, N. V. NMR-based metabolomics. *Drug Chem. Toxicol.* **2002***, 25*, 375-382.

121. Wishart, D. S. Computational strategies for metabolite identification in metabolomics. *Bioanalysis* **2009***, 1*, 1579-1596.

122. Wishart, D. S. Quantitative metabolomics using NMR. *Trends Anal. Chem.* **2008***, 27*, 228-237.

123. Tanaka, K.; Budd, M. A.; Efron, M. L.; Isselbacher, K. J. Isovaleric acidemia: a new genetic defect of leucine metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **1966***, 56*, 236-242.

124. Kuhara, T. Gas chromatographic-mass spectrometric urinary metabolome analysis to study mutations of inborn errors of metabolism. *Mass Spectrom. Rev.* **2005***, 24*, 814-827.

125. Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A. R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* **2006***, 1*, 387-396.

126. Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* **2009***, 37*, 603.

127. Matsumoto, I.; Kuhara, T. A new chemical diagnostic method for inborn errors of metabolism by mass spectrometry-rapid, practical, and simultaneous urinary metabolites analysis. *Mass Spectrom. Rev.* **1996***, 15*, 43-57.

128. Halket, J. M.; Waterman, D.; Przyborowska, A. M.; Patel, R. K.; Fraser, P. D.; Bramley, P. M. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. *J. Exp. Bot.* **2005***, 56*, 219-243.

129. A, J.; Trygg, J.; Gullberg, J.; Johansson, A. I.; Jonsson, P.; Antti, H.; Marklund, S. L.; Moritz, T. Extraction and GC/MS analysis of the human blood plasma metabolome. *Anal. Chem.* **2005***, 77*, 8086-8094.

130. Henion, J.; Skrabalak, D.; Dewey, E.; Maylin, G. Micro LC/MS in drug analysis and metabolism studies. *Drug Metab. Rev.* **1983***, 14*, 961-1003.

131. Rashed, M. S.; Ozand, P. T.; Bucknall, M. P.; Little, D. Diagnosis of inborn errors of metabolism from blood spots by acylcarnitines and amino acids profiling using automated electrospray tandem mass spectrometry. *Pediatr. Res.* **1995***, 38*, 324-331.

132. Dunn, W. B. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Phys. Biol.* **2008***, 5*, 011001-3975/5/1/011001.

133. Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. HMDB 3.0--the human metabolome database in 2013. *Nucleic Acids Res.* **2013***, 41*, 801.

134. Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **2017***, 46*, D608-D617.

135. Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J.

A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* **2009***, 37*, 603.

136. Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; Macinnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. HMDB: the human metabolome database. *Nucleic Acids Res.* **2007***, 35*, 521.

137. Godfrey LV, Glass JB. 2011. Research on Nitrification and Related Processes, Part A. Methods in Enzymology. 486:608 page.

138. Singh, A. Experimental Methodologies for the Characterization of Nanoparticles. *Engineered Nanoparticles* **2016,** 125-170.

139. Beauchemin, D. Inductively Coupled Plasma Mass Spectrometry. *Anal. Chem.* **2010***, 82*, 4786-4810.

140. Hailemariam, D.; Mandal, R.; Saleem, F.; Dunn, S. M.; Wishart, D. S.; Ametaj, B. N. Identification of predictive biomarkers of disease state in transition dairy cows. *J. Dairy Sci.* **2014***, 97*, 2680-2693.

141. LeBlanc, S. J.; Leslie, K. E.; Duffield, T. F. Metabolic predictors of displaced abomasum in dairy cattle. *J. Dairy Sci.* **2005***, 88*, 159-170.

142. Sundekilde, U. K.; Poulsen, N. A.; Larsen, L. B.; Bertram, H. C. Nuclear magnetic resonance metabonomics reveals strong association between milk metabolites and somatic cell count in bovine milk. *J. Dairy Sci.* **2013***, 96*, 290-299.

143. Saleem, F.; Ametaj, B. N.; Bouatra, S.; Mandal, R.; Zebeli, Q.; Dunn, S. M.; Wishart, D. S. A metabolomics approach to uncover the effects of grain diets on rumen health in dairy cows. *J. Dairy Sci.* **2012***, 95*, 6606-6623.

144. Abarghuei, M. J.; Rouzbehan, Y.; Salem, A. Z. M.; Zamiri, M. J. Nitrogen balance, blood metabolites and milk fatty acid composition of dairy cows fed pomegranate-peel extract. *Livestock Science* **2014***, 164*, 72-80.

145. Karisa, B. K.; Thomson, J.; Wang, Z.; Stothard, P.; Moore, S. S.; Plastow, G. S. Candidate genes and single nucleotide polymorphisms associated with variation in residual feed intake in beef cattle. *J. Anim. Sci.* **2013***, 91*, 3502-3513.

146. Widmann, P.; Weikard, R.; Khn, C. Metabolomics: a pathway for improved understanding of genetic modulation of mammalian growth and tissue deposition. *10th World Congress on Genetics Applied to Livestock Production* **2014**.

147. Weikard, R.; Altmaier, E.; Suhre, K.; Weinberger, K. M.; Hammon, H. M.; Albrecht, E.; Setoguchi, K.; Takasuga, A.; Kuhn, C. Metabolomic profiles indicate distinct physiological pathways affected by two loci with major divergent effect on Bos taurus growth and lipid deposition. *Physiol. Genomics* **2010***, 42A*, 79-88.

148. Chapinal, N.; Carson, M. E.; LeBlanc, S. J.; Leslie, K. E.; Godden, S.; Capel, M.; Santos, J. E.; Overton, M. W.; Duffield, T. F. The association of serum metabolites in the transition period with milk production and early-lactation reproductive performance. *J. Dairy Sci.* **2012***, 95*, 1301-1309.

149. Melzer, N.; Wittenburg, D.; Repsilber, D. Integrating milk metabolite profile information for the prediction of traditional milk traits based on SNP information for Holstein cows. *PLoS One* **2013***, 8*, e70256.

150. Melzer, N.; Wittenburg, D.; Hartwig, S.; Jakubowski, S.; Kesting, U.; Willmitzer, L.; Lisec, J.; Reinsch, N.; Repsilber, D. Investigating associations between milk metabolite profiles and milk traits of Holstein cows. *J. Dairy Sci.* **2013***, 96*, 1521-1534.

151. Foroutan, A.; Guo, A. C.; Vazquez-Fresno, R.; Lipfert, M.; Zhang, L.; Zheng, J.; Badran, H.; Budinski, Z.; Mandal, R.; Ametaj, B. N.; Wishart, D. S. Chemical composition of commercial cow's milk. *J. Agric. Food Chem.* **2019***, 67*, 4897-4914.

152. Karisa, B. K.; Thomson, J.; Wang, Z.; Li, C.; Montanholi, Y. R.; Miller, S. P.; Moore, S. S.; Plastow, G. S. Plasma metabolites associated with residual feed intake and other productivity performance traits in beef cattle. *Livestock Science* **2014***, 165*, 200-211.

153. Cameron, N. D. Correlated physiological responses to selection forcarcass lean content in sheep. *Livest. Prod. Sci.* **1992***, 30*, 53-68.

154. Clarke, J. N.; Binnie, D. B.; Jones, K. R.; Mowat, C. M.; Purchas, R. W.; Uljee, A. E. Repeatabilities of blood plasma metabolites and their association with leanness in genotypes showing a wide divergence in carcass composition. *Proc NZ Soc Anim Prod* **1996***, 56*, 180-183.

155. Richardson, E. C.; Herd, R. M.; Archer, J. A.; Arthur, P. F. Metabolic differences in Angus steers divergently selected for residual feed intake. *Aust. J. Exp. Agric.* **2004***, 44*, 441-452.

156. Zambrano, E.; Bautista, C. J.; Deas, M.; Martinez-Samayoa, P. M.; Gonzalez-Zamorano, M.; Ledesma, H.; Morales, J.; Larrea, F.; Nathanielsz, P. W. A low maternal protein diet during pregnancy and lactation has sex- and window of exposure-specific effects on

offspring growth and food intake, glucose metabolism and serum leptin in the rat. *J. Physiol.* **2006***, 571*, 221-230.

157. Gurekian, C. N.; Koski, K. G. Amniotic fluid amino acid concentrations are modified by maternal dietary glucose, gestational age, and fetal growth in rats. *J. Nutr.* **2005***, 135*, 2219-2224.

158. Kwon, H.; Ford, S. P.; Bazer, F. W.; Spencer, T. E.; Nathanielsz, P. W.; Nijland, M. J.; Hess, B. W.; Wu, G. Maternal nutrient restriction reduces concentrations of amino acids and polyamines in ovine maternal and fetal plasma and fetal fluids. *Biol. Reprod.* **2004***, 71*, 901-908.

159. Koski, K. G.; Fergusson, M. A. Amniotic fluid composition responds to changes in maternal dietary carbohydrate and is related to metabolic status in term fetal rats. *J. Nutr.* **1992***, 122*, 385-392.

160. Al-Husseini, W.; Gondro, C.; Quinn, K.; Herd, R. M.; Gibson, J. P.; Chen, Y. Expression of candidate genes for residual feed intake in Angus cattle. *Anim. Genet.* **2014***, 45*, 12-19.

161. Paradis, F.; Wood, K. M.; Swanson, K. C.; Miller, S. P.; McBride, B. W.; Fitzsimmons, C. Maternal nutrient restriction in mid-to-late gestation influences fetal mRNA expression in muscle tissues in beef cattle. *BMC Genomics* **2017***, 18*, 632-5.

162. Paradis, F.; Yue, S.; Grant, J. R.; Stothard, P.; Basarab, J. A.; Fitzsimmons, C. Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. *J. Anim. Sci.* **2015***, 93*, 3331-3341.

163. Koltai, H.; Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic Acids Res.* **2008***, 36*, 2395-2405.

164. Shendure, J. The beginning of the end for microarrays? *Nat. Methods* **2008***, 5*, 585-587.

165. Kukurba, K. R.; Montgomery, S. B. RNA Sequencing and Analysis. *Cold Spring Harb Protoc.* **2015***, 2015*, 951-969.

166. Casneuf, T.; Van de Peer, Y.; Huber, W. In situ analysis of cross-hybridisation on microarrays and the inference of expression correlation. *BMC Bioinformatics* **2007***, 8*, 461-461.

167. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009***, 10*, 57-63.

168. Chu, Y.; Corey, D. R. RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Ther.* **2012***, 22*, 271-274.

169. Adamski, M. G.; Gumann, P.; Baird, A. E. A method for quantitative analysis of standard and high-throughput qPCR expression data based on input sample quantity. *PLoS One* **2014***, 9*, e103917.

170. Baker, M. Digital PCR hits its stride. *Nature Methods* **2012***, 9*, 541-544.

171. Micke, G. C.; Sullivan, T. M.; McMillen, I. C.; Gentili, S.; Perry, V. E. Protein intake during gestation affects postnatal bovine skeletal muscle growth and relative expression of IGF1, IGF1R, IGF2 and IGF2R. *Mol. Cell. Endocrinol.* **2011***, 332*, 234-241.

172. Canani, R. B.; Costanzo, M. D.; Leone, L.; Bedogni, G.; Brambilla, P.; Cianfarani, S.; Nobili, V.; Pietrobelli, A.; Agostoni, C. Epigenetic mechanisms elicited by nutrition in early life. *Nutr. Res. Rev.* **2011***, 24*, 198-205.

173. Moore, L. D.; Le, T.; Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **2013***, 38*, 23-38.

174. Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000***, 403*, 41-45.

175. Romanoski, C. E.; Glass, C. K.; Stunnenberg, H. G.; Wilson, L.; Almouzni, G. Epigenomics: Roadmap for regulation. *Nature* **2015***, 518*, 314-316.

176. Funston, R. N.; Summers, A. F. Epigenetics: setting up lifetime production of beef cows by managing nutrition. *Annu. Rev. Anim. Biosci.* **2013***, 1*, 339-363.

177. Zeisel, S. H. Epigenetic mechanisms for nutrition determinants of later health outcomes. *Am. J. Clin. Nutr.* **2009***, 89*, 1488S-1493S.

178. McKay, J. A.; Mathers, J. C. Diet induced epigenetic changes and their implications for health. *Acta Physiol. (Oxf)* **2011***, 202*, 103-118.

179. Martin, E. M.; Fry, R. C. Environmental Influences on the Epigenome: Exposure-Associated DNA Methylation in Human Populations. *Annu. Rev. Public Health* **2018***, 39*, 309-333.

180. Reik, W.; Walter, J. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2001***, 2*, 21-32.

181. Bartolomei, M. S. Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev.* **2009***, 23*, 2124-2133.

182. Patten, M. M.; Ross, L.; Curley, J. P.; Queller, D. C.; Bonduriansky, R.; Wolf, J. B. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity (Edinb)* **2014***, 113*, 119-128.

183. Ferguson-Smith, A. C. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **2011***, 12*, 565-575.

184. Wood, A. J.; Oakey, R. J. Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genet.* **2006***, 2*, e147.

185. Wagner, C. R. Germ Cells and Epigenetics. *Nature Education* **2010***, 3*, 64.

186. Keil, K. P.; Lein, P. J. DNA methylation: a mechanism linking environmental chemical exposures to risk of autism spectrum disorders? *Environ. Epigenet* **2016***, 2*, 10.1093/eep/dvv012. Epub 2016 Jan 30.

187. Kanherkar, R. R.; Bhatia-Dey, N.; Csoka, A. B. Epigenetics across the human lifespan. *Front. Cell. Dev. Biol.* **2014***, 2*, 49.

188. Carey, M. F.; Peterson, C. L.; Smale, S. T. Chromatin Immunoprecipitation (ChIP). *Cold Spring Harbor Protocols* **2009***, 2009*, pdb.prot5279.

189. Li, Y.; Tollefsbol, T. O. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol. Biol.* **2011***, 791*, 11-21.

190. Olova, N.; Krueger, F.; Andrews, S.; Oxley, D.; Berrens, R. V.; Branco, M. R.; Reik, W. Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biol.* **2018***, 19*, 33-2.

191. Rocha MIP, de Souza MM, Diniz WJS, de Lima AO, Afonso J, Tizioto PC, de Oliveira PSN, Regitano LCA, Koltes JE, Niciura SCM. Liver DNA methylation profile in Nelore cattle extremes for feed efficiency. *I workshop: Pos-Graduacao em Genetica Evolutiva e Biologia Molecular* **2019**, 64.

192. Lan, X.; Cretney, E. C.; Kropp, J.; Khateeb, K.; Berg, M. A.; Penagaricano, F.; Magness, R.; Radunz, A. E.; Khatib, H. Maternal diet during pregnancy induces gene expression and DNA methylation changes in fetal tissues in sheep. *Front. Genet.* **2013***, 4*, 49.

193. Wang, X.; Lan, X.; Radunz, A. E.; Khatib, H. Maternal nutrition during pregnancy is associated with differential expression of imprinted genes and DNA methyltranfereases in muscle of beef cattle offspring. *J. Anim. Sci.* **2015***, 93*, 35-40.

194. Crouse, M. S.; Caton, J. S.; Cushman, R. A.; McLean, K. J.; Dahlen, C. R.; Borowicz, P. P.; Reynolds, L. P.; Ward, A. K. Maternal nutrition during the first 50 days of gestation alters expression of histone and histone modifying genes in bovine fetal liver. *36th International Society of Animal Genetics Conference. Dublin, Ireland* **2017**.

195. Widmann, P.; Reverter, A.; Fortes, M. R.; Weikard, R.; Suhre, K.; Hammon, H.; Albrecht, E.; Kuehn, C. A systems biology approach using metabolomic data reveals genes and pathways interacting to modulate divergent growth in cattle. *BMC Genomics* **2013***, 14*, 798-798.

196. Lu, J.; Boeren, S.; van Hooijdonk, T.; Vervoort, J.; Hettinga, K. Effect of the DGAT1 K232A genotype of dairy cows on the milk metabolome and proteome. *J. Dairy Sci.* **2015***, 98*, 3460-3469.

197. Lee, H. J.; Jung, J. Y.; Oh, Y. K.; Lee, S. S.; Madsen, E. L.; Jeon, C. O. Comparative survey of rumen microbial communities and metabolites across one caprine and three bovine groups, using bar-coded pyrosequencing and (1)H nuclear magnetic resonance spectroscopy. *Appl. Environ. Microbiol.* **2012***, 78*, 5983-5993.

198. Saleem, F.; Bouatra, S.; Guo, A. C.; Psychogios, N.; Mandal, R.; Dunn, S. M.; Ametaj, B. N.; Wishart, D. S. The bovine ruminal fluid metabolome. *Metabolomics* **2013***, 9*, 360-378.

199. O'Callaghan, T. F.; Vazquez-Fresno, R.; Serra-Cayuela, A.; Dong, E.; Mandal, R.; Hennessy, D.; McAuliffe, S.; Dillon, P.; Wishart, D. S.; Stanton, C.; Ross, R. P. Pasture

feeding changes the bovine rumen and milk metabolome. *Metabolites* **2018***, 8*, 10.3390/metabo8020027.

200. Raun, B. M.; Kristensen, N. B. Metabolic effects of feeding ethanol or propanol to postpartum transition Holstein cows. *J. Dairy Sci.* **2011***, 94*, 2566-2580.

201. Jung, Y.; Lee, J.; Kwon, J.; Lee, K. S.; Ryu, D. H.; Hwang, G. S. Discrimination of the geographical origin of beef by (1)H NMR-based metabolomics. *J. Agric. Food Chem.* **2010***, 58*, 10458-10466.

202. Muroya, S.; Oe, M.; Ojima, K.; Watanabe, A. Metabolomic approach to key metabolites characterizing postmortem aged loin muscle of Japanese Black (Wagyu) cattle. *Asian-Australas J. Anim. Sci.* **2019***, 32*, 1172-1185.

203. Kodani, Y.; Miyakawa, T.; Komatsu, T.; Tanokura, M. NMR-based metabolomics for simultaneously evaluating multiple determinants of primary beef quality in Japanese Black cattle. *Sci. Rep.* **2017***, 7*, 1297-1298.

204. Kim, Y. H.; Kemp, R.; Samuelsson, L. M. Effects of dry-aging on meat quality attributes and metabolite profiles of beef loins. *Meat Sci.* **2016***, 111*, 168-176.

205. Miles, E. D.; McBride, B. W.; Jia, Y.; Liao, S. F.; Boling, J. A.; Bridges, P. J.; Matthews, J. C. Glutamine synthetase and alanine transaminase expression are decreased in livers of aged vs. young beef cows and GS can be upregulated by 17beta-estradiol implants. *J. Anim. Sci.* **2015***, 93*, 4500-4509.

206. Brown-Woodman, P. D.; White, I. G. Amino acid composition of semen and the secretions of the male reproductive tract. *Aust. J. Biol. Sci.* **1974***, 27*, 415-422.

207. Sexton, T. J.; Amann, R. P.; Flipse, R. J. Free amino acids and protein in rete testis fluid, vas deferens plasma, accessory sex gland fluid, and seminal plasma of the conscious bull. *J. Dairy Sci.* **1971***, 54*, 412-416.

208. Beauchemin, K. Dietary mitigation of enteric methane from cattle. *Cab Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* **2009***, 4*.

209. CCAC. 1993. Guide to the care and use of experimental animals. Can. Counc. Anim. Care, Ottawa, Ontario, Canada.

210. Mao, F.; Chen, L.; Vinsky, M.; Okine, E.; Wang, Z.; Basarab, J.; Crews, D. H.; Li, C. Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. *J. Anim. Sci.* **2013***, 91*, 2067-2076.

211. Foroutan, A.; Goldansaz, S. A.; Lipfert, M.; Wishart, D. S. Protocols for NMR analysis in livestock metabolomics. *Methods Mol. Biol.* **2019***, 1996*, 311-324.

212. Foroutan, A.; Fitzsimmons, C.; Mandal, R.; Piri-Moghadam, H.; Zheng, J.; Guo, A. C.; Li, C.; Guan, L. L.; Wishart, D. S. The bovine metabolome. *Metabolites* **2020***, 10*, 233.

213. Foroutan, A.; Goldansaz, S. A.; Lipfert, M.; Wishart, D. S. Protocols for NMR analysis in livestock metabolomics. *Methods Mol. Biol.* **2019***, 1996*, 311-324.

214. Magan, J. B.; O'Callaghan, T. F.; Zheng, J.; Zhang, L.; Mandal, R.; Hennessy, D.; Fenelon, M. A.; Wishart, D. S.; Kelly, A. L.; McCarthy, N. A. Impact of bovine diet on metabolomic profile of skim milk and whey protein ingredients. *Metabolites* **2019***, 9*, 10.3390/metabo9120305.

215. Minerbi, A.; Gonzalez, E.; Brereton, N. J. B.; Anjarkouchian, A.; Dewar, K.; Fitzcharles, M. A.; Chevalier, S.; Shir, Y. Altered microbiome composition in individuals with fibromyalgia. *Pain* **2019***, 160*, 2589-2602.

216. Chong, J.; Wishart, D. S.; Xia, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr. Protoc. Bioinformatics* **2019***, 68*, e86.

217. Worley, B.; Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metabolomics* **2013***, 1*, 92-107.

218. Yin, P.; Lehmann, R.; Xu, G. Effects of pre-analytical processes on blood samples used in metabolomics studies. *Anal. Bioanal Chem.* **2015***, 407*, 4879-4892.

219. Artegoitia, V. M.; Middleton, J. L.; Harte, F. M.; Campagna, S. R.; de Veth, M. J. Choline and choline metabolite patterns and associations in blood and milk during lactation in dairy cows. *PLoS One* **2014***, 9*, e103412.

220. Clemmons, B. A.; Mihelic, R. I.; Beckford, R. C.; Powers, J. B.; Melchior, E. A.; McFarlane, Z. D.; Cope, E. R.; Embree, M. M.; Mulliniks, J. T.; Campagna, S. R.; Voy, B. H.; Myer, P. R. Serum metabolites associated with feed efficiency in black angus steers. *Metabolomics* **2017***, 13*, 147.

221. Jorge-Smeding, E.; Renand, G.; Centeno, D.; Pétéra, M.; Durand, S.; Polakof, S.; Cantalapiedra-Hijar, G. Metabolomics reveals changes in urea cycle associated to residual feed intake in growing heifers. *EAAP Scientific Series* **2019**, *138*, 231-232.

222. Basarab, J. A.; Beauchemin, K. A.; Baron, V. S.; Ominski, K. H.; Guan, L. L.; Miller, S. P.; Crowley, J. J. Residual feed intake (RFI): An indirect approach for reducing GHG emissions. *Greenhouse Gases & Animal Agriculture Conference. Dublin, Ireland* **2013**.

223. Tatum, J.D. Animal Age, Physiological Maturity, and Associated Effects on Beef Tenderness. 2011. https://www.beefresearch.org/CMDocs/BeefResearch/PE_White_%20Papers/Animal_Age.pdf. Accessed 10 June 2020.

224. Elolimy, A.; Alharthi, A.; Zeineldin, M.; Parys, C.; Loor, J. J. Residual feed intake divergence during the preweaning period is associated with unique hindgut microbiome and metabolome profiles in neonatal Holstein heifer calves. *J. Anim. Sci. Biotechnol.* **2020***, 11*, 13-x. eCollection 2020.

225. Macdonald, K. A.; Pryce, J. E.; Spelman, R. J.; Davis, S. R.; Wales, W. J.; Waghorn, G. C.; Williams, Y. J.; Marett, L. C.; Hayes, B. J. Holstein-Friesian calves selected for divergence in residual feed intake during growth exhibited significant but reduced residual feed intake divergence in their first lactation. *J. Dairy Sci.* **2014***, 97*, 1427-1435.

226. Beef Cattle Research Council. www.beefresearch.ca/research-topic.cfm/castration-67. Accessed 10 June 2020.

227. Hamilton, T. Beef Bull Fertility. **2015**. http://www.omafra.gov.on.ca/english/livestock/beef/facts/06-015.htm Accessed 10 June 2020.

228. Hall, J.B. The Cow-Calf Manager. **2004**. https://www.sites.ext.vt.edu/newsletter-archive/livestock/aps-04_03/aps-315.html. Accessed 10 June 2020.

229. Yang, Y.; Sadri, H.; Prehn, C.; Adamski, J.; Rehage, J.; Danicke, S.; Saremi, B.; Sauerwein, H. Acylcarnitine profiles in serum and muscle of dairy cows receiving conjugated linoleic acids or a control fat supplement during early lactation. *J. Dairy Sci.* **2019***, 102*, 754-767.

230. Wishart, D. S. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol. Rev.* **2019***, 99*, 1819-1875.

231. Lake, A. D.; Novak, P.; Shipkova, P.; Aranibar, N.; Robertson, D. G.; Reily, M. D.; Lehman-McKeeman, L. D.; Vaillancourt, R. R.; Cherrington, N. J. Branched chain amino acid metabolism profiles in progressive human nonalcoholic fatty liver disease. *Amino Acids* **2015***, 47*, 603-615.

232. Kirchberg, F. F.; Harder, U.; Weber, M.; Grote, V.; Demmelmair, H.; Peissner, W.; Rzehak, P.; Xhonneux, A.; Carlier, C.; Ferre, N.; Escribano, J.; Verduci, E.; Socha, P.; Gruszfeld, D.; Koletzko, B.; Hellmuth, C.; European Childhood Obesity Trial Study Group Dietary protein intake affects amino acid and acylcarnitine metabolism in infants aged 6 months. *J. Clin. Endocrinol. Metab.* **2015***, 100*, 149-158.

233. Bridi, R.; Latini, A.; Braum, C. A.; Zorzi, G. K.; Moacir, W.; Lissi, E.; Dutra-Filho, C. S. Evaluation of the mechanisms involved in leucine-induced oxidative damage in cerebral cortex of young rats. *Free Radic. Res.* **2005***, 39*, 71-79.

234. Gauster, M.; Rechberger, G.; Sovic, A.; Horl, G.; Steyrer, E.; Sattler, W.; Frank, S. Endothelial lipase releases saturated and unsaturated fatty acids of high density lipoprotein phosphatidylcholine. *J. Lipid Res.* **2005***, 46*, 1517-1525.

235. Wong, J. T.; Tran, K.; Pierce, G. N.; Chan, A. C.; O, K.; Choy, P. C. Lysophosphatidylcholine stimulates the release of arachidonic acid in human endothelial cells. *J. Biol. Chem.* **1998***, 273*, 6830-6836.

236. Kim, E. A.; Kim, J. A.; Park, M. H.; Jung, S. C.; Suh, S. H.; Pang, M. G.; Kim, Y. J. Lysophosphatidylcholine induces endothelial cell injury by nitric oxide production through oxidative stress. *J. Matern. Fetal. Neonatal Med.* **2009***, 22*, 325-331.

237. Li, B.; Tian, S.; Liu, X.; He, C.; Ding, Z.; Shan, Y. Sulforaphane protected the injury of human vascular endothelial cell induced by LPC through up-regulating endogenous antioxidants and phase II enzymes. *Food Funct.* **2015***, 6*, 1984-1991.

238. Zou, Y.; Kim, D. H.; Jung, K. J.; Heo, H. S.; Kim, C. H.; Baik, H. S.; Yu, B. P.; Yokozawa, T.; Chung, H. Y. Lysophosphatidylcholine enhances oxidative stress via the 5-lipoxygenase pathway in rat aorta during aging. *Rejuvenation Res.* **2009***, 12*, 15-24.

239. Lehmann, R.; Franken, H.; Dammeier, S.; Rosenbaum, L.; Kantartzis, K.; Peter, A.; Zell, A.; Adam, P.; Li, J.; Xu, G.; Konigsrainer, A.; Machann, J.; Schick, F.; Hrabe de Angelis, M.; Schwab, M.; Staiger, H.; Schleicher, E.; Gastaldelli, A.; Fritsche, A.; Haring, H. U.; Stefan, N. Circulating lysophosphatidylcholines are markers of a metabolically benign nonalcoholic fatty liver. *Diabetes Care* **2013***, 36*, 2331-2338.

240. Stiuso, P.; Scognamiglio, I.; Murolo, M.; Ferranti, P.; De Simone, C.; Rizzo, M. R.; Tuccillo, C.; Caraglia, M.; Loguercio, C.; Federico, A. Serum oxidative stress markers and lipidomic profile to detect NASH patients responsive to an antioxidant treatment: a pilot study. *Oxid Med. Cell. Longev* **2014***, 2014*, 169216.

241. Casal, A.; Garcia-Roche, M.; Navajas, E. A.; Cassina, A.; Carriquiry, M. Differential hepatic oxidative status in steers with divergent residual feed intake phenotype. *Animal* **2020***, 14*, 78-85.

242. Alhasawi, A.; Castonguay, Z.; Appanna, N. D.; Auger, C.; Appanna, V. D. Glycine metabolism and anti-oxidative defence mechanisms in Pseudomonas fluorescens. *Microbiol. Res.* **2015***, 171*, 26-31.

243. Thomas, S. C.; Alhasawi, A.; Auger, C.; Omri, A.; Appanna, V. D. The role of formate in combatting oxidative stress. *Antonie Van Leeuwenhoek* **2016***, 109*, 263-271.

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

245. Harper, A. E.; Miller, R. H.; Block, K. P. Branched-chain amino acid metabolism. *Annu. Rev. Nutr.* **1984***, 4*, 409-454.

246. Monirujjaman, M.; Ferdouse, A. Metabolic and Physiological Roles of Branched-Chain Amino Acids. *Advances in Molecular Biology* **2014***, 2014*, 364976.

247. Freund, H. R.; Hanani, M. The metabolic role of branched-chain amino acids. *Nutrition* **2002***, 18*, 287-288.

248. Kimball, S. R.; Jefferson, L. S. Signaling pathways and molecular mechanisms through which branched-chain amino acids mediate translational control of protein synthesis. *J. Nutr.* **2006***, 136*, 227S-31S.

249. Sun, X.; Zemel, M. B. Leucine and calcium regulate fat metabolism and energy partitioning in murine adipocytes and muscle cells. *Lipids* **2007***, 42*, 297-305.

250. Alberta Agriculture and Forestry Multi-year economic, productive & financial performance of Alberta cow/calf operations (2012-2016). **2017**.

251. Fletcher, T. F.; Weber, A. F. Veterinary development anatomy. Veterinary Embryology. CVM6903. **2013**.

252. Meyer, A. M.; Reed, J. J.; Vonnahme, K. A.; Soto-Navarro, S. A.; Reynolds, L. P.; Ford, S. P.; Hess, B. W.; Caton, J. S. Effects of stage of gestation and nutrient restriction during early to mid-gestation on maternal and fetal visceral organ mass and indices of jejunal growth and vascularity in beef cows. *J. Anim. Sci.* **2010***, 88*, 2410-2424.

253. Ford, S. P.; Hess, B. W.; Schwope, M. M.; Nijland, M. J.; Gilbert, J. S.; Vonnahme, K. A.; Means, W. J.; Han, H.; Nathanielsz, P. W. Maternal undernutrition during early to mid-gestation in the ewe results in altered growth, adiposity, and glucose tolerance in male offspring. *J. Anim. Sci.* **2007***, 85*, 1285-1294.

254. Chen, Y.; Gondro, C.; Quinn, K.; Herd, R. M.; Parnell, P. F.; Vanselow, B. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. *Anim. Genet.* **2011***, 42*, 475-490.

255. Brameld, J. M.; Mostyn, A.; Dandrea, J.; Stephenson, T. J.; Dawson, J. M.; Buttery, P. J.; Symonds, M. E. Maternal nutrition alters the expression of insulin-like growth factors in fetal sheep liver and skeletal muscle. *J. Endocrinol.* **2000***, 167*, 429-437.

256. Li, C.; Schlabritz-Loutsevitch, N. E.; Hubbard, G. B.; Han, V.; Nygard, K.; Cox, L. A.; McDonald, T. J.; Nathanielsz, P. W. Effects of maternal global nutrient restriction on fetal baboon hepatic insulin-like growth factor system genes and gene products. *Endocrinology* **2009***, 150*, 4634-4642.

257. CCAC Guide to the care and use of experimental animals. *Canadian Council on Animal Care. Ottawa, Canada* **1993**.

258. Listrat, A.; Lebret, B.; Louveau, I.; Astruc, T.; Bonnet, M.; Lefaucheur, L.; Picard, B.; Bugeon, J. How muscle structure and composition influence meat and flesh quality. *ScientificWorldJournal* **2016***, 2016*.

259. Realini, C. E.; Venien, A.; Gou, P.; Gatellier, P.; Perez-Juan, M.; Danon, J.; Astruc, T. Characterization of Longissimus thoracis, Semitendinosus and Masseter muscles and relationships with technological quality in pigs. 1. Microscopic analysis of muscles. *Meat Sci.* **2013***, 94*, 408-416.

260. Thapa, B. R.; Walia, A. Liver function tests and their interpretation. *The Indian Journal of Pediatrics* **2007***, 74*, 663-671.

261. Roggeman, S.; de Boeck, G.; De Cock, H.; Blust, R.; Bervoets, L. Accumulation and detoxification of metals and arsenic in tissues of cattle (Bos taurus), and the risks for human consumption. *Sci. Total Environ.* **2014***, 466-467*, 175-184.

262. Uhlen, M.; Fagerberg, L.; Hallstrom, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, A.; Kampf, C.; Sjostedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P. H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Ponten, F. Proteomics. Tissue-based map of the human proteome. *Science* **2015***, 347*, 1260419.

263. Johnson, R. W. Inhibition of growth by pro-inflammatory cytokines: an integrated view. *J. Anim. Sci.* **1997***, 75*, 1244-1255.

264. Spurlock, M. E. Regulation of metabolism and growth during immune challenge: an overview of cytokine function. *J. Anim. Sci.* **1997***, 75*, 1773-1783.

265. Koressaar, T.; Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **2007***, 23*, 1289-1291.

266. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B. C.; Remm, M.; Rozen, S. G. Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **2012***, 40*, e115.

267. Andersen, C. L.; Jensen, J. L.; Orntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited

for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **2004***, 64*, 5245-5250.

268. Zhang, T.; Guan, G.; Chen, T.; Jin, J.; Zhang, L.; Yao, M.; Qi, X.; Zou, J.; Chen, J.; Lu, F.; Chen, X. Methylation of PCDH19 predicts poor prognosis of hepatocellular carcinoma. *Asia Pac. J. Clin. Oncol.* **2018***, 14*, e352-e358.

269. Chen, Y.; Arthur, P. F.; Herd, R. M.; Quinn, K.; Barchia, I. M. Using genes differentially expressed in bulls to classify steers divergently selected for high and low residual feed intake. *Anim. Prod. Sci.* **2012***, 52*, 608-612.

270. de Las Heras-Saldana, S.; Clark, S. A.; Duijvesteijn, N.; Gondro, C.; van der Werf, J H J; Chen, Y. Combining information from genome-wide association and multi-tissue gene expression studies to elucidate factors underlying genetic variation for residual feed intake in Australian Angus cattle. *BMC Genomics* **2019***, 20*, 939-4.

271. Hayes, J. D.; Flanagan, J. U.; Jowsey, I. R. Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **2005***, 45*, 51-88.

272. Radi, R. Oxygen radicals, nitric oxide, and peroxynitrite: Redox pathways in molecular medicine. *Proc. Natl. Acad. Sci. U. S. A.* **2018***, 115*, 5839-5848.

273. Gabory, A.; Ripoche, M. A.; Le Digarcher, A.; Watrin, F.; Ziyyat, A.; Forne, T.; Jammes, H.; Ainscough, J. F.; Surani, M. A.; Journot, L.; Dandolo, L. H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development* **2009***, 136*, 3413-3421.

274. Monnier, P.; Martinet, C.; Pontis, J.; Stancheva, I.; Ait-Si-Ali, S.; Dandolo, L. H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc. Natl. Acad. Sci. U. S. A.* **2013***, 110*, 20693-20698.

275. Bergman, D.; Halje, M.; Nordin, M.; Engstrom, W. Insulin-like growth factor 2 in development and disease: a mini-review. *Gerontology* **2013***, 59*, 240-249.

276. Juszczuk-Kubiak, E.; Starzynski, R. R.; Wicinska, K.; Flisikowski, K. Promoter variant-dependent mRNA expression of the MEF2A in longissimus dorsi muscle in cattle. *DNA Cell Biol.* **2012***, 31*, 1131-1135.

277. Wang, Y. N.; Yang, W. C.; Li, P. W.; Wang, H. B.; Zhang, Y. Y.; Zan, L. S. Myocyte enhancer factor 2A promotes proliferation and its inhibition attenuates myogenic differentiation via myozenin 2 in bovine skeletal muscle myoblast. *PLoS One* **2018***, 13*, e0196255.

278. Matschke, G. H.; Erickson, B. H. Development and radioresponse of the prenatal bovine testis. *Biol. Reprod.* **1969***, 1*, 207-214.

279. Kornfeld, S. Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. *Annu. Rev. Biochem.* **1992***, 61*, 307-330.

280. De Meyts, P. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling. *Diabetologia* **1994***, 37*, 135.

281. Maltin, C. A.; Delday, M. I.; Sinclair, K. D.; Steven, J.; Sneddon, A. A. Impact of manipulations of myogenesis in utero on the performance of adult skeletal muscle. *Reproduction* **2001***, 122*, 359-374.

282. Moresi, V.; Williams, A. H.; Meadows, E.; Flynn, J. M.; Potthoff, M. J.; McAnally, J.; Shelton, J. M.; Backs, J.; Klein, W. H.; Richardson, J. A.; Bassel-Duby, R.; Olson, E. N. Myogenin and class II HDACs control neurogenic muscle atrophy by inducing E3 ubiquitin ligases. *Cell* **2010***, 143*, 35-45.

283. Ba, H. V.; Reddy, B. V.; Hwang, I. Role of calpastatin in the regulation of mRNA expression of calpain, caspase, and heat shock protein systems in bovine muscle satellite cells. *In Vitro Cell. Dev. Biol. Anim.* **2015***, 51*, 447-454.

284. Leal-Gutierrez, J. D.; Elzo, M. A.; Johnson, D. D.; Scheffler, T. L.; Scheffler, J. M.; Mateescu, R. G. Association of mu-calpain and calpastatin polymorphisms with meat tenderness in a Brahman-Angus population. *Front. Genet.* **2018***, 9*, 56.

285. Morgan, J. B.; Wheeler, T. L.; Koohmaraie, M.; Crouse, J. D.; Savell, J. W. Effect of castration on myofibrillar protein turnover, endogenous proteinase activities, and muscle growth in bovine skeletal muscle. *J. Anim. Sci.* **1993***, 71*, 408-414.

286. Nassiri, M.; Tahmoorespur, M.; Javadmanesh, A.; Soltani, M.; Far, S. Calpastatin polymorphism and its association with daily gain in Kurdi sheep. *Iranian Journal of Biotechnology* **2006***, 4*, 188-192.

287. Chung, H.; Davis, M. PCR-RFLP of the ovine calpastatin gene and its association with growth. *Asian Journal of Animal and Veterinary Advances* **2012***, 7*, 641-652.

288. Khan, S.; Riaz, M. N.; Ghaffar, A.; Khan, M. F. U. Calpastatin (CAST) gene polymorphism and its association with average daily weight gain in Balkhi and Kajli sheep and Beetal goat breeds. *Pak. J. Zool.* **2012***, 44*, 377-382.

289. Djureinovic, D.; Fagerberg, L.; Hallstrom, B.; Danielsson, A.; Lindskog, C.; Uhlen, M.; Ponten, F. The human testis-specific proteome defined by transcriptomics and antibody-based profiling. *Mol. Hum. Reprod.* **2014***, 20*, 476-488.

290. Kelly, A. K.; Waters, S. M.; McGee, M.; Browne, J. A.; Magee, D. A.; Kenny, D. A. Expression of key genes of the somatotropic axis in longissimus dorsi muscle of beef heifers phenotypically divergent for residual feed intake. *J. Anim. Sci.* **2013***, 91*, 159-167.

291. Mukiibi, R.; Vinsky, M.; Keogh, K.; Fitzsimmons, C.; Stothard, P.; Waters, S. M.; Li, C. Liver transcriptome profiling of beef steers with divergent growth rate, feed intake, or metabolic body weight phenotypes1. *J. Anim. Sci.* **2019***, 97*, 4386-4404.

292. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **2002***, 16*, 6-21.

293. Hackett, J. A.; Surani, M. A. DNA methylation dynamics during the mammalian life cycle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2013***, 368*, 20110328.

294. Morgan, H. D.; Santos, F.; Green, K.; Dean, W.; Reik, W. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **2005***, 14 Spec No 1*, 47.

295. Sasaki, H.; Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* **2008***, 9*, 129-140.

296. Smith, Z. D.; Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **2013***, 14*, 204-220.

297. Donkin, I.; Barres, R. Sperm epigenetics and influence of environmental factors. *Mol. Metab.* **2018***, 14*, 1-11.

298. Aston, K. I.; Uren, P. J.; Jenkins, T. G.; Horsager, A.; Cairns, B. R.; Smith, A. D.; Carrell, D. T. Aberrant sperm DNA methylation predicts male fertility status and embryo quality. *Fertil. Steril.* **2015***, 104*, 1388-5.

299. Atsem, S.; Reichenbach, J.; Potabattula, R.; Dittrich, M.; Nava, C.; Depienne, C.; Bohm, L.; Rost, S.; Hahn, T.; Schorsch, M.; Haaf, T.; El Hajj, N. Paternal age effects on sperm FOXK1 and KCNA7 methylation and transmission into the next generation. *Hum. Mol. Genet.* **2016***, 25*, 4996-5005.

300. Fang, L.; Zhou, Y.; Liu, S.; Jiang, J.; Bickhart, D. M.; Null, D. J.; Li, B.; Schroeder, S. G.; Rosen, B. D.; Cole, J. B.; Van Tassell, C. P.; Ma, L.; Liu, G. E. Integrating signals from

sperm methylome analysis and genome-wide association study for a better understanding of male fertility in cattle. *Epigenomes* **2019**, *3*, 10.

301. Heijmans, B. T.; Tobi, E. W.; Stein, A. D.; Putter, H.; Blauw, G. J.; Susser, E. S.; Slagboom, P. E.; Lumey, L. H. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17046-17049.
302. Jenkins, T. G.; Aston, K. I.; Meyer, T. D.; Hotaling, J. M.; Shamsi, M. B.; Johnstone, E. B.; Cox, K. J.; Stanford, J. B.; Porucznik, C. A.; Carrell, D. T. Decreased fecundity and sperm DNA methylation patterns. *Fertil. Steril.* **2016**, *105*, 51-3.
303. Jenkins, T. G.; Aston, K. I.; Pflueger, C.; Cairns, B. R.; Carrell, D. T. Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. *PLoS Genet.* **2014**, *10*, e1004458.
304. Tiffon, C. The Impact of Nutrition and Environmental Epigenetics on Human Health and Disease. *Int. J. Mol. Sci.* **2018**, *19*, 10.3390/ijms19113425.
305. Caton, J. S.; Crouse, M. S.; Reynolds, L. P.; Neville, T. L.; Dahlen, C. R.; Ward, A. K.; Swanson, K. C. Maternal nutrition and programming of offspring energy requirements1. *Trans Anim Sci* **2019**, *3*, 976-990.
306. Arthur, P. F.; Archer, J. A.; Johnston, D. J.; Herd, R. M.; Richardson, E. C.; Parnell, P. F. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. *J. Anim. Sci.* **2001**, *79*, 2805-2811.
307. Fox JT. 2004. Characterization of residual feed intake and relationships with performance, carcass and temperament traits in growing calves. Master's thesis, Texas A&M University. URL: http://hdl.handle.net/1969.1/1193.

308. Hafla, A. N.; Lancaster, P. A.; Carstens, G. E.; Forrest, D. W.; Fox, J. T.; Forbes, T. D.; Davis, M. E.; Randel, R. D.; Holloway, J. W. Relationships between feed efficiency, scrotal circumference, and semen quality traits in yearling bulls. *J. Anim. Sci.* **2012***, 90*, 3937-3944.

309. Bruinjé, T. C.; Ponce-Barajas, P.; Dourey, A.; Colazo, M. G.; Caldwell, T.; Wang, Z.; Miller, S. P.; Ambrose, D. J. Morphology, membrane integrity, and mitochondrial function in sperm of crossbred beef bulls selected for residual feed intake. *Can. J. Anim. Sci.* **2019***, 99*, 456-464.

310. Dance, A.; Thundathil, J.; Blondin, P.; Kastelic, J. Enhanced early-life nutrition of Holstein bulls increases sperm production potential without decreasing postpubertal semen quality. *Theriogenology* **2016***, 86*, 687-694.e2.

311. Xi, Y.; Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **2009***, 10*, 232-232.

312. Xiang, H.; Zhu, J.; Chen, Q.; Dai, F.; Li, X.; Li, M.; Zhang, H.; Zhang, G.; Li, D.; Dong, Y.; Zhao, L.; Lin, Y.; Cheng, D.; Yu, J.; Sun, J.; Zhou, X.; Ma, K.; He, Y.; Zhao, Y.; Guo, S.; Ye, M.; Guo, G.; Li, Y.; Li, R.; Zhang, X.; Ma, L.; Kristiansen, K.; Guo, Q.; Jiang, J.; Beck, S.; Xia, Q.; Wang, W.; Wang, J. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat. Biotechnol.* **2010***, 28*, 516-520.

313. Lister, R.; Pelizzola, M.; Dowen, R. H.; Hawkins, R. D.; Hon, G.; Tonti-Filippini, J.; Nery, J. R.; Lee, L.; Ye, Z.; Ngo, Q. M.; Edsall, L.; Antosiewicz-Bourget, J.; Stewart, R.; Ruotti, V.; Millar, A. H.; Thomson, J. A.; Ren, B.; Ecker, J. R. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **2009***, 462*, 315-322.

314. Juhling, F.; Kretzmer, H.; Bernhart, S. H.; Otto, C.; Stadler, P. F.; Hoffmann, S. Metilene: Fast and Sensitive Calling of Differentially Methylated Regions from Bisulfite Sequencing Data. *Genome Res.* **2016**, *26*, 256-262.

315. Olshen, A. B.; Venkatraman, E. S.; Lucito, R.; Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **2004**, *5*, 557-572.

316. Siegmund, D. Boundary crossing probabilities and statistical applications. *The Annals of Statistics* **1986**, *14*.

317. Fasano, G.; Franceschini, A. A multidimensional version of the Kolmogorov–Smirnov test. *Monthly Notices of the Royal Astronomical Society* **1987**, *225*, 155-170.

318. Lawrence, M.; Gentleman, R.; Carey, V. rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* **2009**, *25*, 1841-1842.

319. Zhou, Y.; Connor, E. E.; Bickhart, D. M.; Li, C.; Baldwin, R. L.; Schroeder, S. G.; Rosen, B. D.; Yang, L.; Van Tassell, C. P.; Liu, G. E. Comparative whole genome DNA methylation profiling of cattle sperm and somatic tissues reveals striking hypomethylated patterns in sperm. *Gigascience* **2018**, *7*, 10.1093/gigascience/giy039.

320. Lister, R.; O'Malley, R. C.; Tonti-Filippini, J.; Gregory, B. D.; Berry, C. C.; Millar, A. H.; Ecker, J. R. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **2008**, *133*, 523-536.

321. Datt, C.; Sharma, V.; Dudi, K.; Baban, B.; Sharma, P.; Negesse, T.; Kundu, S. s.; Dutta, M. M.; Gupta, R.; Singh, D. Residual Feed Intake as a Tool for Selecting More Efficient Animals: A Review. *Indian Journal of Animal Nutrition* **2017**, *34*, 238.

322. Radford, E. J.; Ito, M.; Shi, H.; Corish, J. A.; Yamazawa, K.; Isganaitis, E.; Seisenberger, S.; Hore, T. A.; Reik, W.; Erkek, S.; Peters, A H F M; Patti, M. E.; Ferguson-Smith, A. C.

In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science* **2014***, 345*, 1255903.

323. Jimenez-Chillaron, J. C.; Hernandez-Valencia, M.; Reamer, C.; Fisher, S.; Joszi, A.; Hirshman, M.; Oge, A.; Walrond, S.; Przybyla, R.; Boozer, C.; Goodyear, L. J.; Patti, M. E. Beta-cell secretory dysfunction in the pathogenesis of low birth weight-associated diabetes: a murine model. *Diabetes* **2005***, 54*, 702-711.
324. Kropp, J.; Carrillo, J. A.; Namous, H.; Daniels, A.; Salih, S. M.; Song, J.; Khatib, H. Male fertility status is associated with DNA methylation signatures in sperm and transcriptomic profiles of bovine preimplantation embryos. *BMC Genomics* **2017***, 18*, 280-y.
325. Liu, S.; Chen, S.; Cai, W.; Yin, H.; Liu, A.; Li, Y.; Liu, G.; Wang, Y.; Yu, Y.; Zhang, S. Divergence Analyses of Sperm DNA Methylomes between Monozygotic Twin AI Bulls. *Epigenomes* **2019***, 3*, 21.
326. Toschi, P.; Capra, E.; Anzalone, D. A.; Lazzari, B.; Turri, F.; Pizzi, F.; Scapolo, P. A.; Stella, A.; Williams, J. L.; Ajmone Marsan, P.; Loi, P. Maternal peri-conceptional undernourishment perturbs offspring sperm methylome. *Reproduction* **2020***, 159*, 513-523.
327. Stenmark, H.; Olkkonen, V. M. The Rab GTPase family. *Genome Biol.* **2001***, 2*, REVIEWS3007-reviews3007. Epub 2001 Apr 27.
328. Snider, M. D. A role for rab7 GTPase in growth factor-regulated cell nutrition and apoptosis. *Mol. Cell* **2003***, 12*, 796-797.
329. Manning, B. D.; Cantley, L. C. AKT/PKB signaling: navigating downstream. *Cell* **2007***, 129*, 1261-1274.

330. Li, X.; Zhang, X.; Leathers, R.; Makino, A.; Huang, C.; Parsa, P.; Macias, J.; Yuan, J. X.; Jamieson, S. W.; Thistlethwaite, P. A. Notch3 signaling promotes the development of pulmonary arterial hypertension. *Nat. Med.* **2009***, 15*, 1289-1297.

331. Murta, D.; Batista, M.; Trindade, A.; Silva, E.; Henrique, D.; Duarte, A.; Lopes-da-Costa, L. In vivo notch signaling blockade induces abnormal spermatogenesis in the mouse. *PLoS One* **2014***, 9*, e113365.

332. Sagare-Patil, V.; Vernekar, M.; Galvankar, M.; Modi, D. Progesterone utilizes the PI3K-AKT pathway in human spermatozoa to regulate motility and hyperactivation but not acrosome reaction. *Mol. Cell. Endocrinol.* **2013***, 374*, 82-91.

333. de Castro Barbosa, T.; Ingerslev, L. R.; Alm, P. S.; Versteyhe, S.; Massart, J.; Rasmussen, M.; Donkin, I.; Sjogren, R.; Mudry, J. M.; Vetterli, L.; Gupta, S.; Krook, A.; Zierath, J. R.; Barres, R. High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring. *Mol. Metab.* **2015***, 5*, 184-197.

334. Wu, H.; Estill, M. S.; Shershebnev, A.; Suvorov, A.; Krawetz, S. A.; Whitcomb, B. W.; Dinnie, H.; Rahil, T.; Sites, C. K.; Pilsner, J. R. Preconception urinary phthalate concentrations and sperm DNA methylation profiles among men undergoing IVF treatment: a cross-sectional study. *Hum. Reprod.* **2017***, 32*, 2159-2169.

335. Li, C.; Fan, Y.; Li, G.; Xu, X.; Duan, J.; Li, R.; Kang, X.; Ma, X.; Chen, X.; Ke, Y.; Yan, J.; Lian, Y.; Liu, P.; Zhao, Y.; Zhao, H.; Chen, Y.; Yu, Y.; Liu, J. DNA methylation reprogramming of functional elements during mammalian embryonic development. *Cell. Discov.* **2018***, 4*, 41-9. eCollection 2018.

336. Liu, T.; Zhang, L.; Joo, D.; Sun, S. C. NF-kappaB signaling in inflammation. *Signal. Transduct Target Ther.* **2017***, 2*, 10.1038/sigtrans.2017.23. Epub 2017 Jul 14.

337. Skau, C. T.; Plotnikov, S. V.; Doyle, A. D.; Waterman, C. M. Inverted formin 2 in focal adhesions promotes dorsal stress fiber and fibrillar adhesion formation to drive extracellular matrix assembly. *Proc. Natl. Acad. Sci. U. S. A.* **2015***, 112*, 2447.

338. Atzei, P.; Gargan, S.; Curran, N.; Moynagh, P. N. Cactin targets the MHC class III protein IkappaB-like (IkappaBL) and inhibits NF-kappaB and interferon-regulatory factor signaling pathways. *J. Biol. Chem.* **2010***, 285*, 36804-36817.

339. Dai, P.; Jeong, S. Y.; Yu, Y.; Leng, T.; Wu, W.; Xie, L.; Chen, X. Modulation of TLR signaling by multiple MyD88-interacting partners including leucine-rich repeat Fli-I-interacting proteins. *J. Immunol.* **2009***, 182*, 3450-3460.

340. Marangi, G.; Leuzzi, V.; Manti, F.; Lattante, S.; Orteschi, D.; Pecile, V.; Neri, G.; Zollino, M. TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. *Eur. J. Hum. Genet.* **2013***, 21*, 229-232.

341. McLellan, K.; Malik, G.; Davidson, J. P31 Arthritis: a previously unrecognised feature of KBG syndrome. *Rheumatology (Oxford)* **2018***, 57*.

342. Ranganathan, P.; Kattal, N.; Moustafa, M.; Sharma, R.; Thomas, A.; Agarwal, A. Correlation of nuclear factor kappa B (NFKB) with sperm quality and clinical diagnoses in infertile men. *Fertility and Sterility - FERT STERIL* **2002***, 78*.

343. Zhao, Y.; Li, Q.; Yao, C.; Wang, Z.; Zhou, Y.; Wang, Y.; Liu, L.; Wang, Y.; Wang, L.; Qiao, Z. Characterization and quantification of mRNA transcripts in ejaculated spermatozoa of fertile men by serial analysis of gene expression. *Hum. Reprod.* **2006***, 21*, 1583-1590.

344. Abbasi, M.; Williamson, L.; Turek, P. J.; Horsager, A.; Uren, P. J. Semen parameters are associated with alterations in the sperm epigenome of infertile men. *Fertil. Steril.* **2017***, 108*, e17.

345. Nevin, C.; Carroll, M. Sperm DNA methylation, infertility and transgenerational epigenetics. *J Genet Genomic Sci* **2015***, 1*, 004.

346. Verma, A.; Rajput, S.; De, S.; Kumar, R.; Chakravarty, A. K.; Datta, T. K. Genome-wide profiling of sperm DNA methylation in relation to buffalo (Bubalus bubalis) bull fertility. *Theriogenology* **2014***, 82*, 750-9.e1.

347. Camprubi, C.; Salas-Huetos, A.; Aiese-Cigliano, R.; Godo, A.; Pons, M. C.; Castellano, G.; Grossmann, M.; Sanseverino, W.; Martin-Subero, J. I.; Garrido, N.; Blanco, J. Spermatozoa from infertile patients exhibit differences of DNA methylation associated with spermatogenesis-related processes: an array-based analysis. *Reprod. Biomed. Online* **2016***, 33*, 709-719.

348. Okada, H.; Inoue, T.; Kikuta, T.; Kato, N.; Kanno, Y.; Hirosawa, N.; Sakamoto, Y.; Sugaya, T.; Suzuki, H. Poly(ADP-ribose) polymerase-1 enhances transcription of the profibrotic CCN2 gene. *J. Am. Soc. Nephrol.* **2008***, 19*, 933-942.

349. Guo, X.; Day, T. F.; Jiang, X.; Garrett-Beal, L.; Topol, L.; Yang, Y. Wnt/beta-catenin signaling is sufficient and necessary for synovial joint formation. *Genes Dev.* **2004***, 18*, 2404-2417.

350. Hartmann, C.; Tabin, C. J. Wnt-14 plays a pivotal role in inducing synovial joint formation in the developing appendicular skeleton. *Cell* **2001***, 104*, 341-351.

351. Agarwal, A.; Mahfouz, R. Z.; Sharma, R. K.; Sarkar, O.; Mangrola, D.; Mathur, P. P. Potential biological role of poly (ADP-ribose) polymerase (PARP) in male gametes. *Reprod. Biol. Endocrinol.* **2009***, 7*, 143-143.

352. Jha, R.; Agarwal, A.; Mahfouz, R.; Paasch, U.; Grunewald, S.; Sabanegh, E.; Yadav, S. P.; Sharma, R. Determination of Poly (ADP-ribose) polymerase (PARP) homologues in human ejaculated sperm and its correlation with sperm maturation. *Fertil. Steril.* **2009***, 91*, 782-790.

353. Carrell, D. T.; Cairns, B.; Aston, K. I.; Jenkins, T.; Smith, A. D.; Uren, P. J.; Horsager, A. Methods of identifying male fertility status and embryo quality. **2019**.

354. Kerr, G. E.; Young, J. C.; Horvay, K.; Abud, H. E.; Loveland, K. L. Regulated Wnt/beta-catenin signaling sustains adult spermatogenesis in mice. *Biol. Reprod.* **2014***, 90*, 3.

355. MacDonald, B. T.; Tamai, K.; He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev. Cell.* **2009***, 17*, 9-26.

356. Lewis, T. S.; Shapiro, P. S.; Ahn, N. G. Signal transduction through MAP kinase cascades. *Adv. Cancer Res.* **1998***, 74*, 49-139.

357. Whitmarsh, A. J.; Davis, R. J. A central control for cell growth. *Nature* **2000***, 403*, 255-256.

358. Saba-El-Leil, M. K.; Vella, F. D.; Vernay, B.; Voisin, L.; Chen, L.; Labrecque, N.; Ang, S. L.; Meloche, S. An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. *EMBO Rep.* **2003***, 4*, 964-968.

359. Gerrard, D. E.; Okamura, C. S.; Ranalletta, M. A.; Grant, A. L. Developmental expression and location of IGF-I and IGF-II mRNA and protein in skeletal muscle. *J. Anim. Sci.* **1998***, 76*, 1004-1011.

360. Zhang, X.; Wang, W.; Mo, F.; La, Y.; Li, C.; Li, F. Association of residual feed intake with growth and slaughtering performance, blood metabolism, and body composition in growing lambs. *Sci. Rep.* **2017***, 7*, 12681-7.

361. Meyer, A. M.; Vraspir, R. A.; Ellison, M. J.; Cammack, K. M. The relationship of residual feed intake and visceral organ size in growing lambs fed a concentrate- or forage-based diet. *Livestock Science* **2015***, 176*, 85-90.

362. Fitzsimons, C.; Kenny, D. A.; McGee, M. Visceral organ weights, digestion and carcass characteristics of beef bulls differing in residual feed intake offered a high concentrate diet. *Animal* **2014***, 8*, 949-959.

363. Li, M. W.; Mruk, D. D.; Cheng, C. Y. Mitogen-activated protein kinases in male reproductive function. *Trends Mol. Med.* **2009***, 15*, 159-168.

364. Bae, J. W.; Kim, S. H.; Kim, D. H.; Ha, J. J.; Yi, J. K.; Hwang, S.; Ryu, B. Y.; Pang, M. G.; Kwon, W. S. Ras-related proteins (Rab) are key proteins related to male fertility following a unique activation mechanism. *Reprod. Biol.* **2019***, 19*, 356-362.

365. Misteli, T. Beyond the sequence: cellular organization of genome function. *Cell* **2007***, 128*, 787-800.

366. Euskirchen, G.; Auerbach, R. K.; Snyder, M. SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. *J. Biol. Chem.* **2012***, 287*, 30897-30905.

367. Haijes, H. A.; Koster, M. J. E.; Rehmann, H.; Li, D.; Hakonarson, H.; Cappuccio, G.; Hancarova, M.; Lehalle, D.; Reardon, W.; Schaefer, G. B.; Lehman, A.; van de Laar, I M B H; Tesselaar, C. D.; Turner, C.; Goldenberg, A.; Patrier, S.; Thevenon, J.; Pinelli, M.; Brunetti-Pierri, N.; Prchalova, D.; Havlovicova, M.; Vlckova, M.; Sedlacek, Z.; Lopez, E.; Ragoussis, V.; Pagnamenta, A. T.; Kini, U.; Vos, H. R.; van Es, R. M.; van Schaik, R F M A; van Essen, T A J; Kibaek, M.; Taylor, J. C.; Sullivan, J.; Shashi, V.; Petrovski, S.; Fagerberg, C.; Martin, D. M.; van Gassen, K L I; Pfundt, R.; Falk, M. J.; McCormick, E. M.; Timmers, H. T. M.; van Hasselt, P. M. De Novo Heterozygous POLR2A Variants

Cause a Neurodevelopmental Syndrome with Profound Infantile-Onset Hypotonia. *Am. J. Hum. Genet.* **2019***, 105*, 283-301.

368. Ramayo-Caldas, Y.; Ballester, M.; Sanchez, J. P.; Gonzalez-Rodriguez, O.; Revilla, M.; Reyer, H.; Wimmers, K.; Torrallardona, D.; Quintanilla, R. Integrative approach using liver and duodenum RNA-Seq data identifies candidate genes and pathways associated with feed efficiency in pigs. *Sci. Rep.* **2018***, 8*, 558-5.

369. Alkhaled, Y.; Laqqan, M.; Tierling, S.; Lo Porto, C.; Amor, H.; Hammadeh, M. E. Impact of cigarette-smoking on sperm DNA methylation and its effect on sperm parameters. *Andrologia* **2018**.

370. Laqqan, M.; Solomayer, E. F.; Hammadeh, M. Association between alterations in DNA methylation level of spermatozoa at CpGs dinucleotide and male subfertility problems. *Andrologia* **2018***, 50*, 10.1111/and.12832. Epub 2017 Jul 6.

371. Ciraolo, E.; Morello, F.; Hobbs, R. M.; Wolf, F.; Marone, R.; Iezzi, M.; Lu, X.; Mengozzi, G.; Altruda, F.; Sorba, G.; Guan, K.; Pandolfi, P. P.; Wymann, M. P.; Hirsch, E. Essential role of the p110beta subunit of phosphoinositide 3-OH kinase in male fertility. *Mol. Biol. Cell* **2010***, 21*, 704-711.

372. Weber, K. L.; Welly, B. T.; Van Eenennaam, A. L.; Young, A. E.; Porto-Neto, L. R.; Reverter, A.; Rincon, G. Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. *PLoS One* **2016***, 11*, e0152274.

373. Ault-Seay, T. B.; Melchior-Tiffany, E. A.; Clemmons, B. A.; Cordero, J. F.; Bates, G. E.; Flythe, M. D.; Klotz, J. L.; Ji, H.; Goodman, J. P.; McLean, K. J.; Myer, P. R. Rumen and Serum Metabolomes in Response to Endophyte-Infected Tall Fescue Seed and Isoflavone Supplementation in Beef Steers. *Toxins (Basel)* **2020***, 12*, 10.3390/toxins12120744.

374. Wu, H.; Naya, F. J.; McKinsey, T. A.; Mercer, B.; Shelton, J. M.; Chin, E. R.; Simard, A. R.; Michel, R. N.; Bassel-Duby, R.; Olson, E. N.; Williams, R. S. MEF2 responds to multiple calcium-regulated signals in the control of skeletal muscle fiber type. *EMBO J.* **2000***, 19*, 1963-1973.

375. Mears, D. Regulation of insulin secretion in islets of Langerhans by Ca(2+)channels. *J. Membr. Biol.* **2004***, 200*, 57-66.

376. Everman, S.; Meyer, C.; Tran, L.; Hoffman, N.; Carroll, C. C.; Dedmon, W. L.; Katsanos, C. S. Insulin does not stimulate muscle protein synthesis during increased plasma branched-chain amino acids alone but still decreases whole body proteolysis in humans. *Am. J. Physiol. Endocrinol. Metab.* **2016***, 311*, E671-E677.

377. Nascimento, C. F.; Branco, R. H.; Bonilha, S. F.; Cyrillo, J. N.; Negrao, J. A.; Mercadante, M. E. Residual feed intake and blood variables in young Nellore cattle. *J. Anim. Sci.* **2015***, 93*, 1318-1326.

378. Karlstaedt, A.; Vasquez, H.; Taegtmeyer, H. Abstract 384: Leucine, serine and methionine differentially affect protein synthesis in rat heart. *Circulation Res.* **2015***, 117*, A384.

379. El-Hafidi, M.; Franco, M.; Ramirez, A. R.; Sosa, J. S.; Flores, J. A. P.; Acosta, O. L.; Salgado, M. C.; Cardoso-Saldana, G. Glycine Increases Insulin Sensitivity and Glutathione Biosynthesis and Protects against Oxidative Stress in a Model of Sucrose-Induced Insulin Resistance. *Oxid Med. Cell. Longev* **2018***, 2018*, 2101562.

380. Amelio, I.; Cutruzzola, F.; Antonov, A.; Agostini, M.; Melino, G. Serine and glycine metabolism in cancer. *Trends Biochem. Sci.* **2014***, 39*, 191-198.

381. Siddik, M. A. B.; Shin, A. C. Recent Progress on Branched-Chain Amino Acids in Obesity, Diabetes, and Beyond. *Endocrinol. Metab. (Seoul)* **2019***, 34*, 234-246.