Institute of Animal Science

The serum proteome and metabolome of dairy cows clustered for body fat content from a large dairy herd

Dissertation

For the Degree

Doctor of Philosophy

(Ph.D.)

Faculty of Agriculture

Rheinische Friedrich-Wilhelms-Universität Bonn

by

Ruben Riosa

from Trieste

Bonn, 2022

Referentin:	Prof. Dr. Dr. Helga Sauerwein
Korreferenten:	Dr. Richard Burchmore
	Prof. Dr. Kathryn Elmer
Tag der mündlichen Prüfung:	30.05.2022

Dissertation from the European Joint Doctorate in Molecular Animal Nutrition (MANNA). This project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-MSCA- ITN-2017- EJD: Marie Skłodowska-Curie Innovative Training Networks (European Joint Doctorate) – Grant agreement nº: 765423 – MANNA

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn.

The serum proteome and metabolome of dairy cows clustered for body fat content from a large dairy herd

Ruben Riosa

MSc Science and Technology of Animal Production

BSc Breeding and Animal Health

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

University of Glasgow

Institute of Biodiversity, Animal Health and Comparative Medicine
College of Medical, Veterinary & Life Sciences

2022

Abstract

The transition from late pregnancy to early lactation implies metabolic and endocrine changes for accomplishing the adaptation to the rapid increase of milk production. Voluntary feed intake can usually not cover the energy and nutrient requirements in the first weeks of lactation, and dairy cows thus need to mobilize body reserves, mainly from adipose tissue. The extent of this mobilization that can be assessed by recording backfat thickness (BFT), varies between animals but is commonly more pronounced in cows that are over-conditioned at calving. Over-conditioned cows are at greater risk for developing metabolic disorders, such as ketosis, and thus for compromised welfare and performance than cows of normal or lean body condition. Making use of a large dataset including also health records from a herd with 1,709 multiparous Holstein cows, the objectives of this thesis work were (1) to characterize the variation in pre-calving back fat thickness (BFT) and the subsequent BFT loss during early lactation, and to relate it to milk production, health condition, and selected blood variables, (2) to perform an untargeted metabolomics analysis for comparing the metabolome in blood serum of selected subgroups differing in body condition loss, health status and in dietary methionine (Met) supply, and (3), to undertake proteome analyses in other subgroups of animals that were either lean or over-conditioned before calving but were otherwise not differing in health status and Met supply. Animals from which serum samples and BFT records were available both at day 25 *ante partum* (ap) and day 30 *post partum* (pp) were selected (n =713) and subjected to K-means cluster analyses. Five clusters were obtained each considering the BFT-*ap* and the difference between BFT-ap and BFT pp (ΔBFT). The clusters were validated and the serum samples analysed for non-esterified fatty acids (NEFA), ß-hydroxybutyrate (BHB), for two adipokines, i.e., adiponectin and leptin, and for one inflammation marker (Haptoglobin). In confirmation of the literature, cows in the clusters with greater ΔBFT underwent more intense lipolysis and ketogenesis than cows with smaller ΔBFT. Cows categorized as very fat ap had lesser milk yields than other clusters. No differences in the serum metabolome at day 30 pp were detectable in cows with different ΔBFT, health status, and Met supply (n = 184). Even though the subset was further limited to fat versus lean cows (n = 30 in total) that were all healthy and did not receive supplemental Met for the proteome analysis, no differences were observed between the two groups. The findings about the classical variables recorded were largely confirmatory whereas the multivariate results from metabolomics and proteomics could not further extend the current knowledge about the relationship between body condition, fat mobilization, and metabolism.

Kurzfassung

Der Übergang von der späten Gravidität zur frühen Laktation bringt metabolische und endokrine Veränderungen mit sich, um die Anpassung an den schnellen Anstieg der Milchproduktion zu bewältigen. Die freiwillige Futteraufnahme kann den Energie- und Nährstoffbedarf in den ersten Wochen der Laktation in der Regel nicht decken, so dass Milchkühe Körperreserven, v.a. Fett, mobilisieren müssen. Das Ausmaß dieser Mobilisierung, das durch die Erfassung der Rückfettdicke (BFT) beurteilt werden kann, variiert tierindividuell, ist aber bei Kühen, die beim Abkalben überkonditioniert sind, meist stärker ausgeprägt. Überkonditionierte Kühe haben ein höheres Risiko für die Entwicklung von Stoffwechselstörungen wie Ketose und damit für eine Beeinträchtigung des Wohlbefindens und der Leistung als Kühe mit normaler oder magerer Körperkondition. Unter Verwendung eines großen Datensatzes, der auch Gesundheitsdaten aus einer Herde mit 1.709 pluriparen Holstein-Kühen enthielt, waren die Ziele dieser Arbeit: (1) die BFT-Variation vor dem Abkalben (*ante partum*, ap) und des anschließenden BFT-Verlustes während der frühen Laktation in ihrer Beziehung zu Milchleistung, Gesundheitszustand und ausgewählten Blutvariablen zu charakterisieren, (2) das Metabolom im Blutserum ausgewählter Untergruppen, die sich in Bezug auf den Verlust der Körperkondition, den Gesundheitszustand und die Methionin (Met)-Zufuhr mit der Nahrung unterscheiden, mittels einer *untargeted* Metabolomics-Analyse zu vergleichen, (3) auch das Proteom in anderen Untergruppen von Tieren, die vor dem Kalben entweder mager oder überkonditioniert waren, sich aber ansonsten nicht in Bezug auf den Gesundheitszustand und die Met-Zufuhr unterschieden, zu untersuchen. Es wurden zunächst Tiere ausgewählt, von denen Serumproben und BFT-Aufzeichnungen sowohl am Tag 25 ap als auch am Tag 30 *post partum* (pp) verfügbar waren, und einer *K-Means-Cluster*-Analyse unterzogen. Sowohl für den ap BFT-Wert als auch für die Differenz zwischen BFT-ap und BFT-pp (ΔBFT) wurden jeweils fünf Cluster erhalten. Die Cluster wurden validiert und die Serumproben auf nicht veresterte Fettsäuren (NEFA), ß-Hydroxybutyrat (BHB), auf zwei Adipokine (Adiponectin und Leptin) und auf einen Entzündungsmarker (Haptoglobin) untersucht. Wie in der Literatur beschrieben, durchliefen die Kühe in den Clustern mit größerem ΔBFT eine intensivere Lipolyse und Ketogenese als Kühe mit kleinerem ΔBFT. Kühe, die als sehr fett eingestuft wurden, hatten eine geringere Milchleistung als andere Cluster. Bei Kühen mit unterschiedlichem ΔBFT, Gesundheitszustand und Met-Versorgung (n = 184) waren keine Unterschiede im Serum-Metabolom am Tag 30 pp nachweisbar. Auch die für die Proteom-Analysen vorgenommene Reduzierung der Untergruppen auf fette und magere Kühe (insgesamt n = 30), die alle gesund waren und kein zusätzliches Met erhielten, wurden keine Unterschiede zwischen den beiden Gruppen festgestellt. Die Ergebnisse zu den erfassten klassischen Variablen waren weitgehend bestätigend, während die multivariaten Ergebnisse aus Metabolomics und Proteomics unter den gewählten Bedingungen das derzeitige Verständnis der Beziehungen zwischen Körperkondition, Fettmobilisierung und Stoffwechsel nicht weiter vertiefen konnten.

Table of contents

List of Tables

List of figures

List of abbreviations

2DE	Two-dimensional electrophoresis
2DE-DIGE	2-dimensional difference gel electrophoresis
AA	Amino acids
AGC	Automatic gain control
ap	ante partum
APP	Acute-phase proteins
AST	Aspartate amino transferase
BC	Body condition
BCS	Body condition score
BFT	Back fat thickness
BHB	ß-hydroxybutyrate
CE/MS	Capillary electrophoresis–mass spectrometry
CI	Conception interval
CLA	Conjugated linoleic acids
DDA	Data dependent acquisition
DIA	Data independent acquisition
DIM	Days in milk
DMI	Dry matter intake
ECM	Energy corrected milk
FA	Fatty acids
FCM	Fat corrected milk
FDR	False discovery rates
GC/MS	Gas chromatography–mass spectrometry
GGT	γ-glutamyl transferase

GN	Gain
HCD	Higher energy collision-induced dissociation
Hp	Haptoglobin
HPLC	High Performance Liquid Chromatography
HYK	Hyperketonaemia
IGF-1	Insulin-like growth factor-1
JF	Just fat
KNN	K-nearest neighbour algorithm
LC/MS	Liquid chromatography–mass spectrometry
LEN	Lean
Met	Methionine
ML	Moderate loss
MP	Metabolizable protein
MRM	Multiple-reaction monitoring
mzXML	Extensible markup language for mass-spectrometry based data
NEB	Negative energy balance
NEFA	Non-esterified fatty acids
NIESI	Negative ion electrospray ionisation
NL	No loss
NMR	Nuclear magnetic resonance
NOR	Normal
OC	Over-conditioned
PC	Principal component
PCA	Principal component analysis
peNDF	physically effective NDF

PIESI	Positive ion electrospray ionisation
PiMP	The Polyomics integrated metabolomics pipeline
PLS-DA	Partial least squares-discriminant analysis
PMN	Polymorphonuclear neutrophils
pp	post partum
RPM	Rumen protected methionine
RSD	Relative standard deviation
SDC	Sodium deoxycholate
SF	Slightly fat
SL	Small loss
TFA	Trifluoroacetic acid
TMR	Total mixed ration
UPLC	Ultra-performance liquid chromatography–mass spectrometry
VF	Very fat
VLDL	Very Low Density Lipoprotein
VSN	Variance stabilization and calibration package
XML	eXtensible Markup Language

Acknowledgements

I would like to thank my supervisor, Prof. Dr. Dr. Helga Sauerwein, for being a great mentor. Your skills, your knowledge, your training, and your passion for the research world have been fundamental to shaping me into a better researcher, and I will take all these teachings with me for the rest of my career.

Secondly, I would like to thank Dr. Richard Burchmore and Dr. Mark McLaughlin for your precious advice and support during the period at the University of Glasgow, but also for your constant support during my PhD.

Thank you also to Dr. Morteza Hosseini Ghaffari, Dr. Gavin Blackburn, Dr. Marc Sylvester, and Dr. Farhad Sakheri. Working with you has been extremely motivating and your experience has been invaluable. I would also like to thank Thomas Blees, Barbara Heitkönig, Dr. Ana Monteiro, and all the trainees who worked on my project for having been always so helpful around the lab; but more in general I would like to thank both the members (Ph.D. students, Post-docs, researchers…) of the Institute of Physiology at the University of Bonn and the ones of the School of Veterinary Medicine at the University of Glasgow with whom I have shared many moments.

Thirdly, I would like to thank the whole MANNA consortium. Starting from our project coordinator, Prof. David Eckersall, who was always present and available to answer all questions; Dave Iglesias, our project manager, who was always keen in organising and managing all the activities concerning the MANNA; and finally thank you also to all my fellow Early Stage Researchers: I've met ten incredible researchers that I can now call friends.

Fourthly, I would like to thank Evonik Operations GmbH, represented by Dr. Claudia Parys and Dr. Jessie Guyader for the possibility they have given me to work with them and understand the industrial point of view of a scientific research. And I would like to also thank Prof. Marc Drillich and Dr. Micheal Iwersen from the Faculty of Veterinary Medicine of the University of Vienna for having supported our project by sharing the samples they collected in their study, giving complete access to all the data they have collected, as well as constantly supporting our research.

I would like to thank Lucie for her precious support during the sometimes difficult phases of doing a Ph.D.

Finally, I am extremely grateful for the unconditional love and support that I've received from my wonderful family throughout my studies. I will be always grateful for the opportunities you have given me and that permitted me to become the person I am today.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: **Ruben Riosa**

Chapter 1 - Introduction

1.1 The dairy cow's transition from pregnancy to lactation and the interrelationship with body condition

The so-called transition period in dairy cows was defined as the time from 3 weeks before to 3 weeks after parturition (Drackley, 1999). Especially in early lactation, dairy cows are typically in a status of negative energy balance (NEB) because voluntary feed intake is insufficient to meet the increased nutrient requirements for milk synthesis (Drackley, 1999). This reduction in feed intake may result in a less efficient adaptation to lactation. The transition period is overall characterized by major changes in the digestive, metabolic, and endocrine system with actions that must be operative at the time of calving (Figure 1). The NEB caused by the reduced dry matter intake and the increased energy demand after parturition leads to an increased lipolysis and gluconeogenesis and is also associated with a higher ketogenic status (i.e., higher level of BHB in the blood), which causes an increased inflammatory status post-partum. Due to this status, the high cellular metabolism and upregulated immune gene expression increase even more the energy requirements of the animals in a phase in which there is a reduced dry matter intake. The failure of one or more of these systems may lead to metabolic diseases, for example, ketosis, but also to infectious diseases such as mastitis, and metritis due to a compromised immune defence.

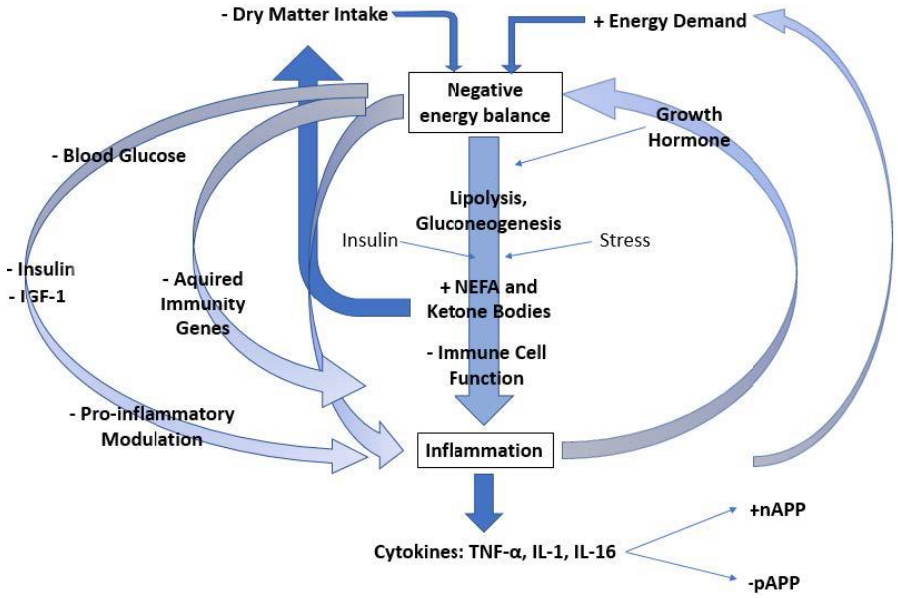


Figure 1. *Major interactions between the immune, endocrine and metabolic systems in dairy cows during the transition period. Inflammation post-partum, intense cellular metabolism and upregulated immune gene expression increase energy requirements when dry matter (from Esposito et al., 2014)*

All these diseases compromise the cow's well-being and may lead to culling or death (McGuffey, 2017). Another important aspect to consider for the health of the animals is the body condition around calving. In optimal circumstances, dairy cows are capable of efficiently using body fat to produce high milk yields already in the first weeks of lactation. However, cows with a high body condition score (BCS; ranging from 1, very thin cow, to 5, over-conditioned cow) around calving are at greater risk of developing metabolic disorders, such as ketosis but also impaired fertility (Gillund et al., 2001; Bernabucci et al., 2005; Roche et al., 2009; Rathbun et al., 2017). Moreover, it is important to underline that not only the BCS around calving is a key element, but also the extent of BCS loss during the transition from pregnancy to lactation has important implications in the metabolic health of dairy cows (Rathbun et al., 2017). A loss of one unit of BCS has been shown to double the risk of ketosis (Duffield et al., 1998; Rathbun et al., 2017). Moreover, an extensive loss of the body's energy reserves is also associated with impaired immune function, which may increase the risk for infectious diseases such as endometritis

(Torres et al., 2020). Body condition around calving is thus considered one of the most relevant determinants of BCS loss: the greater the BCS is at calving, the more body reserves will be mobilized, leading to an increased risk of metabolic disorders (Roche et al., 2007b). It is thus recommended to avoid over-conditioning in animals in the weeks before parturition, and this can be assessed using ultrasound measurements of the backfat thickness (BFT; Schröder and Staufenbiel, 2006), which are more precise and reproducible compared to a visual assessment of the BCS.

The major problem when studying and evaluating how dairy cows accumulate body reserves and how these reserves are mobilized is that they vary among different animals, even if they are fed and managed in the same way, and the reasons for this variation are still mostly unknown. In previous studies using proteomics, it was observed that over-conditioned cows around calving exhibited changes in pathways related to the acute inflammatory response and in the regulation of complement and coagulation cascades (Ghaffari et al., 2020b). Similarly, using a targeted metabolomics approach, Ghaffari et al. (2020a) identified divergent metabotypes even within well-characterized phenotypes of over-conditioned dairy cows at calving with a severe subsequent loss of BC.

With this background, it is important to understand what the requirements of the animals are for supporting their productivity, and to avoid losses in performance or even, in serious situations, the loss of the animal, all of which would decrease the dairy farmer's income. Particular focus should thus be on the transition period, during which dairy cattle start developing a NEB and tend to reduce their dry matter intake (DMI) at a time when nutrient intake should optimally support foetal growth and milk production, respectively (Bell et al., 2000).

Drackley (1999) reported that cows with health disorders around calving produced 7.2 kg less milk per day during the first 20 d post-partum period as compared to healthy cows. Concerning the relation between BCS at calving and milk yield, there are conflicting

reports in the literature. As comprehensively reviewed by Roche et al. (2009), there is no clear linear relationship between BCS and milk yield, however, over-conditioned cows tend to produce less milk.

In case of prolonged metabolic disorders, the most common problems in dairy cows are: parturient paresis (milk fever or hypocalcaemia), fatty liver, ketosis, and laminitis (McGuffey, 2017). In the next sub-sections, the classical variables in blood that should be considered to control the health of the dairy cows will be highlighted, and some strategies to improve animal health around calving will be outlined. In this regard, seeking further improvements in nutritional management is essential: supporting the animal in this phase, and fully understand the metabolic processes which take place during the transition period is the key to have healthy and high-producing dairy cows at the farm level.

1.1.1 Measurements in blood plasma or serum for evaluating metabolic health in dairy cows

To evaluate the health and the metabolic response in dairy cows during the transition period, and to identify potential differences between animals with different BCS before calving or with a different BCS loss, a common approach is to test the concentration of various parameters in the circulation. In relation to individual analytes the spectrum reaches from enzyme activities such as aspartate amino transferase (AST) or γ-glutamyl transferase (GGT), metabolites (e.g., non-esterified fatty acids (NEFA), ß-hydroxybutyrate (BHB), calcium, and glucose), metabolic hormones (e.g., insulin, leptin, adiponectin, growth hormone and insulin-like growth factor-1 (IGF-1)), and markers of oxidative status and inflammation. Besides these “classical” variables that are largely evaluated by univariate statistical models, the use of OMICs technologies for assessing the entire metabolome or the proteome is gaining increasing interest.

In the next paragraphs, some of the "classical variables" that were also assessed in the experimental approaches within this thesis will be introduced in more detail. The selection of the following variables has been made to evaluate the lipolytic and ketogenic status of the animals (NEFA, BHB), the relation between body condition and body condition loss in relation to the endocrine activity of the adipose tissue (leptin and adiponectin), and finally the inflammation status of the cows (haptoglobin).

Non-esterified fatty acids (NEFA)

The NEB is related to mobilization of body reserves, mainly fat. In consequence of increasing lipolysis, more NEFA are released into the bloodstream (Li et al., 2016) and then reach the liver. There NEFA can be *a)* oxidized for ATP production, *b)* deposited as triglycerides in the hepatocytes (which can lead to a fatty liver syndrome), *c)* exported from the liver via lipoproteins, or *d)* can be partially oxidized to ketone bodies (Reynolds et al., 2003). The evaluation of the NEFA concentrations is particularly important to evaluate dry cows' management and nutrition in dairy herds. The concentration in serum can be used as an indicator of the energy balance of the animals: the greater the concentration of NEFA is, associated with a reduced oxidation, the more severe are the negative energy balance and the rate of lipolysis.

ß-hydroxybutyrate (BHB)

As highlighted in the previous paragraph, elevated concentrations of NEFA are associated with increased lipolysis. When the animals are not entirely able to deal with the increased level of NEFA, ketogenesis will increase, and may result in pathological increases in the circulating concentration of ketone bodies.

The BHB is considered as a ketone body (together with acetone and acetoacetate), albeit it is, strictly speaking, not a ketone body since it is lacking a keto group. Ketogenesis increases when the NEFA, released from adipose tissue, cannot be completely oxidized (Ingvartsen, 2006). This is particularly pronounced when lipolysis is excessive (Benedet et al., 2019). Elevated concentrations of ketone bodies in the blood are defined by common thresholds as hyperketonaemia (HYK), which can be subclinical (BHB $\geq$ 1.2 μmol/L) or clinical (BHB > 1.4 μmol/L) (Duffield et al., 2009). The HYK may negatively affect milk production, health, and immunity. The BHB is the most commonly used analyte to diagnose HYK as it is the most stable and most abundant ketone body in blood and also in milk (Duffield et al., 2009). Body condition is also a factor that is related to the concentration of BHB in blood. Gärtner et al. (2019) observed that over-conditioned cows *ante partum* (*ap*) and cows with an extensive mobilization had greater circulating BHB concentrations than cows at normal BCS, underlining that over-conditioned animals are less able to overcome the metabolic challenges of the transition period, thus entering a metabolic imbalance.

Apart from metabolites such as NEFA and BHB, the assessment of the concentration of various metabolic hormones can also provide insights into the metabolic status. The adipokines, i.e. hormones produced by the adipose tissue that can act in an endocrine but also in an autocrine/paracrine manner, are of particular interest. Adipokines may also indicate the metabolic changes around calving; in dairy cows, the best studied adipokines are leptin and adiponectin.

Leptin

Leptin plays an important role in glucose and lipid metabolism, by stimulating lipolysis, inhibiting lipogenesis, and increasing fatty acid oxidation (Block et al., 2003;

Chilliard et al., 2005). Leptin is also involved in regulating metabolic functions connected with insulin sensitivity, and inflammation (Ahima and Lazar, 2008; Galic et al., 2010). The circulating leptin concentrations are positively related to body fat and increasing leptin concentrations have a negative feedback on feed intake (Morrison et al., 2001).

The concentration of leptin in dairy cows varies depending on the stage of the lactation cycle. Specifically, leptinaemia is higher one month before calving and then decreases between the last 4 weeks before calving, reaching the lowest concentration around the first week *post partum* (*pp*), before increasing again during the subsequent weeks (Chilliard et al., 2005).

It is also important to underline that the BCS of the animals is highly related to the leptin concentration, meaning that fatter animals ante-partum have greater concentrations of plasma leptin (Meikle et al., 2004).

Adiponectin

Amongst the adipokines, adiponectin is one of the most abundant (Chilliard et al., 2005). In contrast to leptin, adiponectin tends to have lower concentrations in over-conditioned individuals (Kadowaki and Yamauchi, 2005). Adiponectin inhibits lipolysis in the adipose tissue and decreases insulin resistance by increasing fatty acid oxidation (Yamauchi et al., 2001).

As mentioned before, in dairy cows, adiponectin might be an indicator of the metabolic changes around calving. Overall, its concentration gradually decreased before parturition, reaching a nadir at calving, and then slowly increasing thereafter (Singh et al., 2014; Sauerwein and Häußler, 2016). Thus, the reduced concentration of adiponectin around calving is necessary to promote peripheral insulin resistance, which leads to better nutrient partitioning to the mammary gland in which the uptake of glucose is insulin-independent (Bell, 1995).

Haptoglobin (Hp)

During the transition period, dairy cows normally display some signs of inflammatory response related to the metabolic switch between pregnancy and lactation, even without showing signs of infection (Sordillo et al., 2009). However, excessive inflammatory reactions require substantial amounts of energy, mainly glucose and can thus lead to metabolic diseases. Therefore sick animals should be identified early to ensure a fast treatment and thus increase welfare at the herd level (Barragan et al., 2018). For this reason, the use of biomarkers of inflammation, such as the acute phase protein haptoglobin, was suggested to evaluate the level of inflammation in dairy cows (Huzzey et al., 2009). Haptoglobin is an acute-phase protein and a non-specific marker of inflammation (Ceciliani et al., 2012) and can be useful to detect common transition disorders at an early stage (Huzzey and Overton, 2013). It is mainly released by hepatocytes during an inflammatory event in response to proinflammatory cytokines (Ametaj et al., 2005). The Hp concentration is higher with increasing severity of the inflammatory responses. In dairy cows, Hp is used as a marker of inflammation that could be used in clinical diagnosis, by monitoring the health throughout the lactation period (Ceciliani et al., 2012). Particularly relevant is the association of a high concentration of Hp with an increased incidence of calving difficulties, retained placenta, metritis, and endometritis, as well as a decrease in milk production in dairy herds (Huzzey et al., 2009; Shin et al., 2018). Thus, screening dairy cows for Hp in research trials can be an important tool for assessing the health status.

1.1.2 Strategies to improve health around calving

The majority of health problems in dairy cows happens during the periparturient period; these problems result from an unsuccessful metabolic adaptation to the changes from gestation to lactation (Ingvartsen, 2006) and comprise mainly fatty liver, ketosis,

rumen acidosis, and displaced abomasum. However, these problems are multifactorial diseases that are inter-related and are also associated with inadequate feeding and/or management. In the next paragraphs, an overview of the main strategies to reduce the incidence of these pathologies is provided and represented in Figure 2.

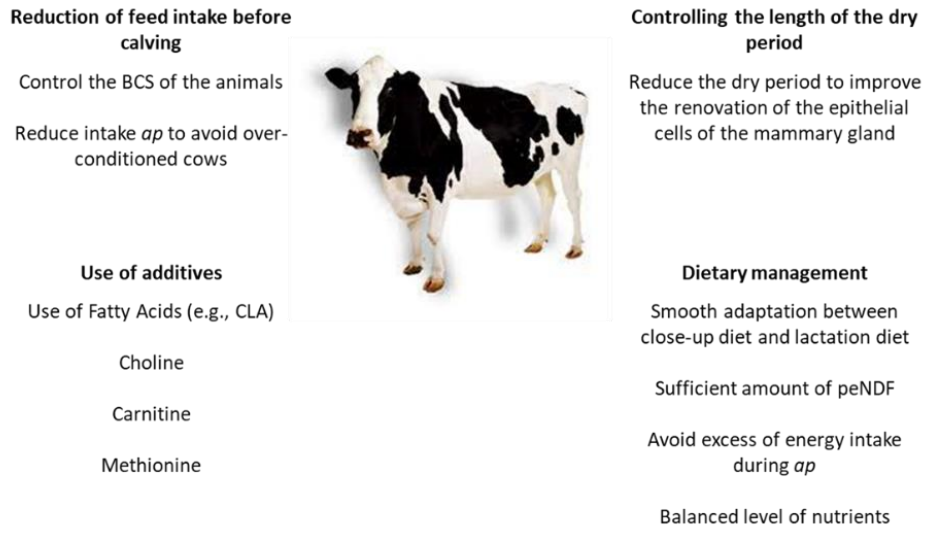


Figure 2. *Summary of the most common strategies to improve dairy cows' health around calving*

Reduction of feed intake before calving

Over-conditioned cows have a greater risk of developing pathologies around calving (Grummer, 1993). As widely reported in the literature, there is a positive correlation between the *ap* body weight and the amount of mobilization, meaning that fatter animals would mobilize more body reserves, thus increasing the risk of metabolic diseases. The body condition of dairy cows should then be at its optimum, i.e., between 3.0 and 3.25 (Roche et al., 2009) on a 1–5 scale (1: very thin, 5: very fat; (Edmonson et al., 1989). A BCS in this range is considered as optimal for adapting to possible dietary inadequacies without having to excessively mobilize body reserves but also supporting milk production (Ingvartsen, 2006). For this reason, it is important to focus on providing adequate diets to achieve the targeted BCS at calving.

However, feed intake restriction before calving might be beneficial, as it increased DMI after parturition as well as milk yield: when limiting feed intake in dairy cows during the prepartum period to 80% of the predicted requirements the peripartal NEFA curve was demonstrated to peak at lesser concentrations than in animals fed *ad libitum* (Overton and Waldron, 2004), and also to increase insulin sensitivity (Holtenius et al., 2003).

Controlling the length of the dry period

The dry period is the time of roughly 6 to 8 weeks before calving in which cows are not milked, and during which cows have a resting period before giving birth necessary to maximize milk production in the subsequent lactation (Kok et al., 2017). Not milking the animals during this time permits the mammary gland to renew its epithelial cells at a faster rate compared to when it would be producing milk (Capuco et al., 1997), resulting in a higher number of renewed cells at calving, which can explain the subsequent peak of production after calving.

New management strategies that were proposed to reduce the NEB in early lactation include a shortening or even the omission of the dry period. Thereby the peak of milk yield can be reduced or postponed to later stages of lactation when voluntary feed intake has sufficiently increased to meet the metabolic requirements (Grummer and Rastani, 2004). By the overall reduction of NEB in the first weeks after calving, the incidence of metabolic disturbances may also be reduced (Kok et al., 2019).

Dietary management

The metabolic changes that dairy cows undergo during the onset of lactation are also strictly connected to an increase in their nutrient requirements, which is mostly related

to the need of the animal to synthesize milk lactose, fat, and protein in the mammary gland (Bell, 1995): a cow producing 30 kg of milk at day 4 after calving, requires 2.7, 4.5, and 2.0 fold greater amounts of glucose, fatty acid, and amino acids (AA) as compared to the *ap* situation. This increase is not followed by the same increase in DMI after calving, in contrast, the DMI in the first week pp can even be reduced to 30-50% of the values of late pregnancy (Roche et al., 2007a). For this reason, the cows need to mobilize body fat reserves, as well as protein and calcium, and also increase hepatic gluconeogenesis. This in turn may increase the risk of metabolic diseases as described in the previous paragraphs. Thus, apart from the management strategy related to the length of the dry period, feeding is of central importance for the metabolic adaptation of dairy cows during the transition period (Roche et al., 2013b).

A proper feeding strategy needs first of all to guarantee the integrity and the health of the rumen, which plays a key role in the dairy cows' overall health, e.g. it modulates the host's immune responses (Zebeli and Metzler-Zebeli, 2012). A smooth adaptation between the close-up and start of lactation is fundamental (Drackley and Cardoso, 2014), and a sufficient amount of physically effective NDF (peNDF) in the diet is needed, to prevent pH depression and thus acidosis. The increased need for energy requires more energy-dense rations but feeding more concentrates without allowing for adequate adaptation of the ruminal wall and the microbiota, can lead to the respective health disturbance and thus exacerbate the situation. Besides diet composition, the actual feeding management is also extremely important in this phase to protect rumen health (Zebeli et al., 2015). For example, giving too large meals in a very short amount of time could predispose dairy cows to rumen disorders because of the decrease of pH in the rumen itself (Krause and Oetzel, 2006). This is because of the reduced salivary secretion when cows ingest a large meal in a short time, which results in the decrease of the capacity of the rumen to adapt that results in the reduction of the rumen pH (Beauchemin et al., 2008).

Attention should be especially paid to avoid an excess of energy intake during the prepartum phase, as the cows should not increase their BCS, otherwise it could become a risk factor as described in the previous paragraphs. However, increasing the amount of concentrate in the close-up is important to allow the rumen epithelium to adapt and at the same time, it will prepare the microbiota for grain-rich diets, typical in the fresh period right after parturition (Drackley and Cardoso, 2014).

Another important aspect to consider is the level of other nutrients, such as metabolizable protein, vitamins, and minerals (Drackley and Cardoso, 2014). The most relevant problem when analysing the mineral balance can be identified as hypocalcaemia. However, this is not in focus of this thesis and will thus not be addressed in detail.

During the transition period, lipid mobilization and the NEB were primarily investigated; however, protein mobilization is another fundamental aspect that should be considered (Bell et al., 2000). The dietary levels of protein in the prepartum diet influence the tissue protein accretion. For this reason, increasing the level of the metabolizable protein (MP) supply can be beneficial as it can increase the reserves in maternal tissues that can be used by the cows after calving (Cardoso et al., 2021). The MP is defined as the true protein which is absorbed by the intestine and that results from both the microbial protein and the proteins bypassing the rumen. Increasing the MP supply is very important as it can also increase milk production, as reported by Larsen et al. (2014).

Use of additives to improve dietary management

Apart from feeding animals with a balanced diet, the use of feed additives can also support dairy cows in adapting to the changing nutrient demands which happen during the transition period.

The dietary use of fatty acids is obvious when aiming at increasing the energy content of the diet, however, the portion of fat that can be included in a ruminant's diet is limited,

otherwise it would compromise ruminal fermentation. Specific fatty acids, i.e., conjugated linoleic acids (CLA), provide the possibility of improving animal health around the transition period (Esposito et al., 2013). Feeding CLA in a form protected from rumen digestion reduced the milk fat (Castañeda-Gutiérrez et al., 2007) and thus decreased the energy content of the milk. Milk fat synthesis normally requires more than the 50% of the energy needed for total milk production (Bauman et al., 2011). Thus, a reduction of milk fat synthesis can reduce the energy demand for milk production (De Veth et al., 2004) and may finally alleviate metabolic stress. Other papers corroborated this notion demonstrating that CLA can improve animal health and performance by reducing NEFA and BHB, and increasing the DMI (Baumgard et al., 2001; Bernal-Santos et al., 2003; Pappritz et al., 2011).

Choline is another important supplement commonly used in dairy cows` nutrition. Choline is a trimethylated quaternary amine which is involved in the transport of fat from the liver, it provides methyl groups for transmethylation reactions, and is required in the synthesis of phosphatidylcholines in cell membranes (Hartwell et al., 2000; Zahra et al., 2006). In dairy cows' nutrition, it has been demonstrated that supplementing rumen-protected choline can increase milk fat synthesis (Erdman et al., 1984) and can be used to treat cows affected by fatty liver after parturition (Cooke et al., 2007).

Carnitine is another important supplement which is essential for the mitochondrial *β*-oxidation of long-chain fatty acids (Carlson et al., 2007). A supplementation of L-carnitine can improve the utilisation of NEFA, and it is associated as a limiting factor for fatty acid metabolism (LaCount et al., 1995). Supplementing the diet of dairy cows with L-carnitine lead to a reduction of triglyceride accumulation in the liver (Carlson et al., 2006), and an increased level of BHB in plasma (Pirestani and Aghakhani, 2018), which can indicate a modulation in the lipid and energy metabolism (Meyer et al., 2020).

The importance of the dietary level of protein has been mentioned before; however, more specifically, the supply with AA, not just crude protein, is crucial for performance and health of dairy cows. The ways of achieving this supply also have environmental consequences: feeding high levels of protein will not necessarily improve the supply of AA since the AA composition, not just the amount of protein, determine the nutritional value. Protein of high biological quality is degraded in the rumen and the supply of essential AA is thus not ensured. Moreover, high quality protein is expensive, not only in terms of economic value but also when considering land use and therefore it should rather go to non-ruminants or directly to human consumption; when feeding low quality protein to ruminants, the AA supply is compromised and will result in increased nitrogen excretion but not accretion in animal products, thus increasing the environmental impact (Lee et al., 2012). For these reasons, supplementation of individual AA is an effective alternative. Feeding dairy cows with a balanced AA supply was demonstrated to increase milk production (Gidlund et al., 2015). Insufficient supply of essential AA may interrupt protein synthesis (Kim and Lee, 2021). Amongst the essential AA in dairy cows, methionine (Met) has been widely studied, and is particularly relevant in this thesis. The animal trial from which data and samples were used, was designed to test the effects of this AA in terms of milk production, health, and body condition of dairy cows.

Methionine (Met)

Methionine is one of the essential AA, and it is formed by an asymmetrical molecule that can occur as L-methionine (*laevus*, left*)* or as D-methionine (*dexter*, right). The DL-Met is a mixture of the two forms, typically represented in equal amounts, producing a racemic mixture. It is non-polar, aliphatic, and contains sulphur (as does cysteine).

The supplementation of Met to dairy cows during the transition period is not a new feeding strategy. The first paper on feeding a rumen-protected Met was published by Broderick et al. (1970; cited by Schwab and Broderick, 2017) who fed Met to early- to mid-lactation Holstein cows. The rumen-protection is important to circumvent the degradation by the rumen microorganisms, otherwise, the half-life in the rumen would be approximately 2.4 h only (Emery, 1971). Today, there are different types of rumen-protected Met with different core-protection approaches (i.e. enzyme resistant, pH-sensitive, lipid film). Furthermore, another important aspect that must be considered while supplementing AA is the ratio between Lysine and Met. It is known that maintaining a ratio of Lys to Met of 3:1 will maximise the metabolic availability and thus help meeting the productive requirements of dairy cows (McGuffey, 2017).

Methionine is typically considered as the first-limiting AA for lactating cows (NRC, 2001) and among the multiple biological functions besides milk protein synthesis for which Met availability is important, the most relevant functions, especially during the transition period, include lipoprotein synthesis in the liver, synthesis of antioxidant proteins, and synthesis of immune-related proteins (Zhou et al., 2016b; Batistel et al., 2017), as summarised in Figure 3. Moreover, Met plays a critical role in the synthesis of carnitine, which is involved in lipid metabolism (Chandler and White, 2017).

More general, the effects of the dietary Met supplementation in dairy cows can be seen at different levels: in terms of milk yield, in terms of health and fertility, and finally at the metabolic level, where it plays a crucial role as methyl-donor.

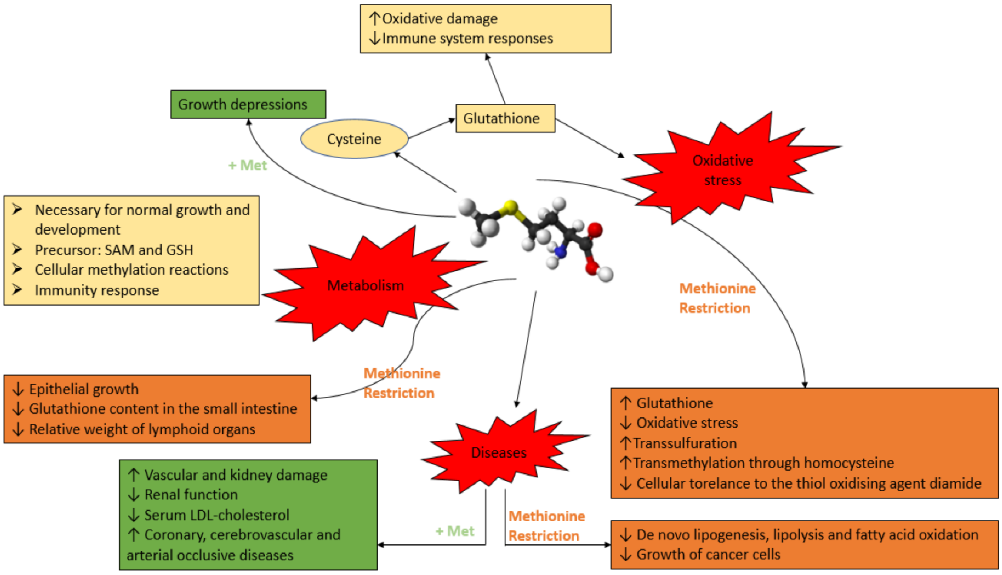


Figure 3. *Main outcome of Met supplementation and of Met restriction on metabolism, oxidative stress, and diseases in mammals (from* Martínez et al. 2017)

Effect of methionine on milk yield in dairy cows

The potential effects of Met on milk yield were also investigated, and are also summarized on Table 1. Batistel et al., (2017) observed that supplementing Met at 0.09 and 0.10% (~20 g/d) of the DM in the *ap* and *pp* periods, respectively (to ensure a ratio of Lys to Met in metabolizable protein close to 2.8:1) with an ethyl-cellulose rumen-protected Met (RPM), Mepron (Evonik Nutrition and Care GmbH, Hanau-Wolfgang, Germany), to multiparous cows (n = 30) increased their DMI (+1.65 kg/d), milk yield (+4.1 kg/d), fat yield (+0.175 kg/d), milk protein yield (+0.2 kg/d), milk lactose yield (+0.25 kg/d), 3.5% FCM (+4.3 kg/d) and ECM (+4.4 kg/d) as compared to control cows not receiving the Met supplement (n = 30). These positive results on milk production were likely driven by the increased DMI and the improved liver function. These results were partially confirmed by Osorio et al. (2013) who supplemented 15 multiparous cows with MetaSmart (MS) (0.19% of DM; ~25 g/d), another coated Methionine composed of small beads of Met with a specific pH-sensitive coating that protects the AA during its passage through the animal's rumen (Adisseo Inc., Antony, France). Similarly, they also supplemented 18 multiparous

cows (n = 24 in the control group) with Smartamine MTM (SM) (0.07% of DMI; ~ 8 g/d), an ester of the analog Met designed to optimize AA balancing in dairy cow rations (Adisseo Inc.). They found, when Lys is adequate to achieve an approximately 2.9:1 Lys:Met ratio, an improvement in milk production (+2.4 kg/d for MS and +4.3 kg/d for SM) which was related to an increased in voluntary DMI. This could be explained by the Met characteristic of being a lipotropic agent which can help to clear lipid accumulation from the liver by stimulating, at least in part, hepatic Very Low Density Lipoprotein (VLDL) formation, and thus supporting lipid export into the circulation and providing energy to the organism (Osorio et al., 2013).

In contrast, Chen et al. (2011) tested 70 lactating cows (50 multiparous and 20 primiparous) and did not find any effect on DMI, milk yield, or 3.5% fat corrected milk (FCM) in animals receiving RPM (Smartamine MTM, Adisseo, Alpharetta, GA, USA) at 2% of DMI. However, the authors found differences in ECM yield, which was higher in Met supplemented cows (+3.1 kg/d). Moreover, Piepenbrink et al. (2004) tested 48 dairy cows, feeding 16 animals with 0.09% or 0.18% of DM of a Met analog (DL-2-hydroxy-4-methylthiobutanoic acid), but did not detect increases in milk yield or fat yields. Similarly, Preynat et al. (2009) fed 24 cows with 9 and 18 g of Mepron-85/d (Degussa AG, 6 Hanau, Germany) pre- and post-calving, respectively, and did not find any effect on milk yield, fat yields, or protein yields.

These studies, selected as the most representative of the works already published, are partly contradictory. However, the reported results were likely not unequivocal because different types of Met were used, and the conditions and the experimental designs were not the same in the various studies.

Table 1. *Summary of the effect of a Methionine supplementation on milk yields, health, and reproduction according to the most representative works already published.*

Reference	**Type of supplementation**	**Observed effects** (compared to a control group of animals)
Effects on milk yields		
Batistel et al. 2017	~20 g/d of Meprom[1] both *ap* and *pp*	DMI (+1.65 kg/d), milk yield (+4.1 kg/d), fat yield (+0.175 kg/d), milk protein yield (+0.2 kg/d), milk lactose yield (+0.25 kg/d), 3.5% FCM (+4.3 kg/d) and ECM (+4.4 kg/d)
Osorio et al. 2013	~25 g/d of MetaSmart[2] both *ap* and *pp*	Milk yield (+2.4 kg/d), increased voluntary DMI
Osorio et al. 2013	~8 g/d of Smartamine[3] both *ap* and *pp*	Milk yield (+4.3 kg/d), increased voluntary DMI
Chen et al. 2011	2 % of DM of Smartamine both *ap* and *pp*	No differences on DMI, milk yield, or 3.5% fat corrected milk
Piepenbrick et al. 2004	0.09% or 0.18% of DM of a Met analog (DL-2-hydroxy-4-methylthiobutanoic acid) both *ap* and *pp*	No differences on milk or fat yields
Preynat et al. 2009	9 and 18 g of Mepron-85/d both *ap* and *pp*	No differences on milk, fat, or protein yields
Effects on health		
Zhou et al. 2016	0.08% of DM of Smartamine both *ap* and *pp*	Better immune response: greater phagocytosis and increased oxidative burst Greater levels of albumin Greater levels of interleuikin-6
Vailati-Riboni et al. 2017	0.08% of DM of Smartamine both *ap* and *pp*	Reduced pro-inflammatory hyper- response
Effects on fertility		
Ardalan and Rezayazdi 2020	18 g/d of Smartamine both *ap* and *pp*	Reduced calving to conception interval (-0.8 services) and days open (-26.8 days)
Suess et al. 2019	27 g/d of Meprom *pp*	No differences in reproductive performances

Effects of methionine on health and fertility in dairy cows

As mentioned above, during the transition period, the immunometabolic status of dairy cows is altered due to changes in liver function, inflammation, and oxidative stress. Feeding management may improve the physiological state and the immune status of the animals (Zhou et al., 2016). One strategy could be supplementing Met directly, or its derivate metabolites (e.g. glutathione, taurine, polyamines), which are well-known immune-nutrients in non-ruminants, that may support the immune functions by increasing

their activity (Vailati-Riboni et al., 2017). The effect of Met on health and fertility in ruminants was tested only recently, but the idea is that Met supplementation will exert beneficial effects on dairy cows' innate immune system. The work of Zhou et al. (2016) was supported by the finding that supplementing Met, Smartamine MTM (Adisseo NA, Alpharetta, GA, USA) at a rate of 0.08% of DM, 21 cows (81 in total) resulted in a better immune response in terms of greater phagocytosis and oxidative burst capabilities upon pathogen challenges. Met-supplemented cows had greater blood albumin concentrations compared with cows fed without Met supplemented, which might indicate that Met-supplemented cows were in better health since albumin is considered as a negative acute-phase protein (Eckersall, 1995). Surprisingly, interleukin-6, which acts mostly as a pro-inflammatory cytokine, was also higher in Met-supplemented cows, which led also to a greater blood neutrophil oxidative burst. These results were confirmed also by Vailati-Riboni et al. (2017), who found that a Met supplementation, Smartamine MTM (Adisseo NA, Alpharetta, GA, USA) at a rate of 0.08% of DM, to 30 cows (60 in total) was able to lessen the pro-inflammatory hyper-response which cows undergo around parturition.

Reproductive performance is another major aspect that must be taken into consideration for a good management. Reproduction is highly influenced by nutrition and management during the transition period. Few works have studied the effect of Met on cows' fertility, but Ardalan and Rezayazdi (2010) found that feeding 18 g/d of rumen-protected methionine (RPM), Smartamine MTM (Adisseo, Antony, France) to 10 cows (n=40 in total: n=10 received rumen-protected choline, n=10 both Met and choline, and n=10 were the control group) from 4 weeks *ap* onwards reduced services per conception (-0.8 services) and days open of lactating dairy cows (-26.8 d) but did not have any effect on days to first oestrus and the number of pregnant cows. In general, this might be explained by the longer period of a negative protein balance in higher-producing cows, putting them at greater risk of reproductive problems, because of the tight connection between protein

metabolism and fertility (Ardalan and Rezayazdi, 2010). Therefore, feeding RPM could help the cows' metabolism to adapt to the challenges in the transition period and thus lead to improved fertility. On the other side, Süss et al. (2019) recently reported from their study with 1,863 cows in total, that supplementation half of these cows with approximately 27 g/d rumen-protected Met (Mepron®, Evonik Nutrition & Care GmbH, Hanau, Germany) did not lead to any differences between the 2 groups with regard to reproductive performance, evaluated as the percentage of dairy cows which were successfully inseminated, either at the first insemination, either in the following ones. The study of Süss et al. (2019) is also the basis of the present thesis. The contradictory results might be in general explained by different farm management practices, or by the different time and duration of the Met supplementation, which could influence the effectiveness of the treatment itself.

1.2 Evaluation of multivariate data by use of machine learning

For the planned use of OMICs analyses in the large herd trial mentioned above (Süss et al. 2019), the number of samples should be reduced for reasons of costs, and also for selecting samples from animals that were most representative for the status to be tested (i.e., Met supplementation or not, healthy or not, divergent in body condition). In view of the great number of animals and samples, such a selection only "by eye" is certainly not state of the art or acceptably objective. Machine learning is a field of computer science in which the machine is trained to complete some specific tasks. It has developed from the area of research of computational learning and pattern recognition in the area of artificial intelligence (Dhall et al., 2020). Training is thus on some data sets, and then, various algorithms are applied so that predictions can be made to learn more from the given data sets (Das et al., 2015), by thus establishing patterns and relevant relationships between parameters.

Within the machine learning analyses, there are two main basic approaches: supervised and unsupervised learning. The main difference between the two is that one uses labeled data to help predict outcomes (supervised), whilst the other does not (unsupervised), as summarised in Table 2.

Table 2. *Comparison between supervised and unsupervised learning (from Dhall et al., 2020)*

Parameters	**Supervised learning**	**Unsupervised learning**
Definition	Supervised learning is defined as that type of machine learning in which there is input and output variables and an algorithm helps to understand the mapping function from input to output	Unsupervised learning is a type of machine learning that includes only input and no output variables
Type of data	The data in supervised learning is labelled, and it predicts the output from the given input data	The data in unsupervised learning is unlabelled, and prediction is made by inheriting structure from the given input data
Accuracy	Supervised learning provides more reliable and accurate results	Unsupervised learning produces average results as compared to the supervised learning
Complexity	The computational complexity in the case of supervised learning is too complex	The computational complexity in the case of unsupervised learning is less

1.2.1 Supervised learning

Supervised algorithms need to use labelled data, which are split into two different parts: a testing data set and a training data set (Kotsiantis et al., 2006). Moreover, the trained data set has some output that has to be predicted, and thus the task is to make these

algorithms learn from similar patterns obtained from the various test on the training data set to then apply them to the testing data set to predict the output results.

In general, supervised learning can be divided into two major types when analysing data:

- Classification: known also as logistic regression, it is used to classify between two or more classes (Dey, 2016).
- Regression: it is used to study the relationship between dependent and independent variables (Dhall et al., 2020).

1.2.2 Unsupervised learning

Unsupervised algorithms do not require labelled data as the algorithm is based solely on the input data. Thus, the training data set is used both for creating and training the model itself; on the other side, the testing data set is used for helping the prediction of the correct values (Dey, 2016). The final prediction is thus based on the outcome of the previous test, and it learns from the previously used features.

In general, unsupervised learning can be divided into two major types when analysing data:

- Clustering: it is used for grouping unlabelled data based on their similarities or differences. The most common algorithm is the k-means, as it creates k-distinct groups (or clusters) of similar data (Dey, 2016).
- Dimensionality reduction: it is used when the number of features (or dimensions) in the dataset is too big, and thus, needs to be reduced. For example, the Principal Component Analysis (PCA) is an algorithm that converts two-dimensional data into one-dimensional data. These techniques are used to make complex data more accessible and fast to analyse (Dey, 2016).

1.2.3 Using machine learning in animal sciences

In the field of animal science, especially when working on big datasets, such as the one derived from complex system biology analyses, these machine learning approaches are starting to be widely used as they can provide actionable knowledge when working on large data sets while aiming to improve metabolic profiling research (Ghaffari et al., 2019). Also when studying animal behaviour, machine learning approaches can solve tasks that would be otherwise time-consuming or too complex to analyse, such as classifying species, individuals, or even behaviours in complex data sets (Valletta et al., 2017).

Machine learning is a powerful tool as it can give a hypothesis-free approach which can be needed in complex data sets in which the relationship of the measured variables is unknown, and thus, where many classical statistical models would fail to detect any differences; a machine learning approach might open new ideas and hypotheses to describe a data set variation (Valletta et al., 2017). As briefly explained in the previous sections, many are the available algorithms that can be used, ranging from regression, classification, to clustering and dimensionality reduction, just to mention the most common ones. With the increased number of data collected, and the increase in power of computers, these methods are of rapidly increasing interest, especially with the development of easy-to-use packages (Valletta et al., 2017).

1.3 Metabolomics

Metabolomics is the comprehensive study of all the metabolites in an organism which are influenced by genetic and environmental factors (Singh et al., 2019). Metabolites are defined as small molecules (whose mass ranges from 50 to 1500 Daltons (Da), such as sugars, lipids, AA, and fatty acids. Thus, they represent all the molecules that have disparate physical properties and which represent the intermediate or final products of complex biological interactions (Clish, 2015).

Hence, metabolomics is the study that allows a complete and very precise description of the metabolome, defined as the complete set of metabolites found within a cell, tissue, or biological sample in a specific time point which are necessary for the growth, maintenance or normal function in a specific physiological state (Oliver et al., 1998; Harrigan and Goodacre, 2003; Singh et al., 2019).

Metabolomics can then be divided into two major approaches: targeted and untargeted (Roberts et al., 2012). Each of the methods have their own advantages and disadvantages (Table 3), which will be discussed in the next paragraphs.

Table 3. *Untargeted versus Targeted Metabolomics Studies. Adapted from* Schrimpe-Rutledge et al. (2016)

Metabolomics studies	
Targeted Methods	**Untargeted Methods**
For validation/quantification	For discovery
Hypothesis-driven	Hypothesis-generating
Subset analyses	Broad analyses
Correlated to reference standards	Correlated to databases and/or libraries
Identification known (quantitative analyses)	Identification not known (qualitative analyses)
Can permit an absolute quantification	Relative quantification

1.3.1 Targeted metabolomics

Targeted (or validation-based) metabolomics uses multiple-reaction monitoring (MRM) mode in order to determine a predefined set of metabolites (Zhou et al., 2016), which in turn permits obtaining both high sensitivity and selectivity. Thereby targets of interest can be identified by excluding background signals of the matrix (Cao et al., 2020).

These advantages qualify targeted metabolomics as the gold standard for the quantification of metabolites (Zhou et al., 2016). This method is thus ideal when trying to measure well-defined groups of metabolites and to confirm a hypothesis (Schrimpe-Rutledge et al., 2016), i.e., quantitative comparison of two groups from which you should expect metabolic differences at a specific time-points.

1.3.2 Untargeted metabolomics

In contrast to the previous method, untargeted (or discovery-based) metabolomics focuses on the overall detection associated with a relative quantitation of the metabolome (Schrimpe-Rutledge et al., 2016).

Due to its comprehensive nature, this method is usually coupled with chemometric techniques, such as multivariate analyses, in order to reduce the large datasets that are generated into a smaller set of manageable signals (Roberts et al., 2012). It is important to highlight that these signals then need to be identified using analytical chemistry, as the method itself does not identify the individual metabolites during the analysis itself. For this reason, untargeted analyses are mostly used when novel target are to be discovered, since the targeted method is restricted by definition to a previously selected asset of targets (Roberts et al., 2012).

1.3.3 Methods of analyses

From an analytical perspective, there also are several methods to pursue metabolomics studies. Presently, the most popular methods are gas chromatography–mass spectrometry (GC/MS), liquid chromatography–mass spectrometry (LC/MS), and capillary electrophoresis–mass spectrometry (CE/MS) (Singh et al., 2019). Within the LC/MS techniques, two of the most relevant are the High Performance Liquid Chromatography (HPLC), ultra-high-performance liquid chromatography–mass spectrometry (UHPLC)

(Scalbert et al., 2009; Singh et al., 2019), but will not be further discussed as they are not in focus of this thesis.

Another key technique to perform metabolomics analysis is represented by the nuclear magnetic resonance (NMR), one of the most widely used technique of metabolomics which provides direct information on the chemical structure. It is a non-destructive, relatively fast, and highly reproducible high throughput analytical platform that requires a reduced sample preparation (Sotelo and Slupsky, 2013).

In order to select the appropriate methodology, dynamic range, accuracy, precision, selectivity, coverage, detection limit, and price, should be considered.

As an example, the classical workflow for processing metabolomics (Figure 4) data sets would include:

- Sample collection and preparation;
- Data acquisition using one of the methods highlighted above (e.g., NMR, LC/MS, GC/MS);
- Spectral processing: which includes all the data processing relative to noise filtering, peak alignment, normalization, etc.;
- Data analysis (via supervised or unsupervised analysis) and metabolites identification;
- Biological interpretation (e.g., pathway analysis)

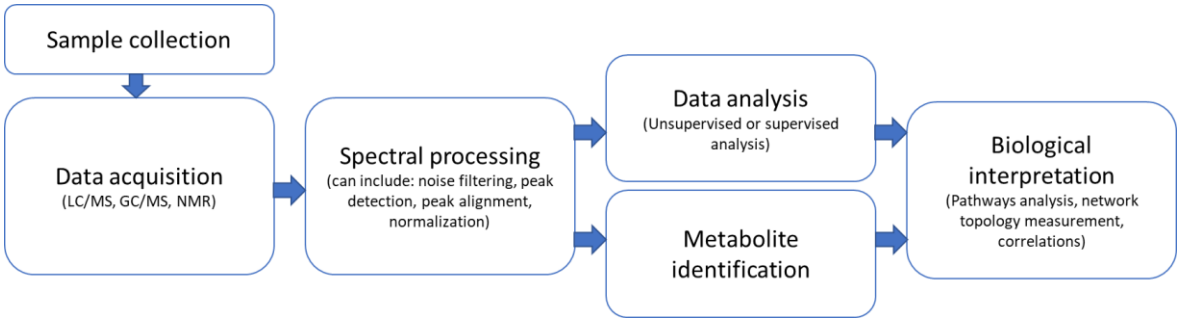


Figure 4. *Example of a typical workflows for untargeted metabolomics analyses*

1.3.4 Application of metabolomics in animal science

Metabolomics analysis, by using advanced analytical chemistry techniques, permits measurement of large numbers of metabolites quickly and comprehensively in the targeted individual (Goldansaz et al., 2017), permitting many scientists to obtain a more complete picture of system-wide metabolism and biology. Metabolomics thus became an increasingly popular "omics" approach which assists in an efficient phenotypic characterization of microorganisms, plants, but also humans and animals (Goldansaz et al., 2017). Especially in animal science, metabolomics can now be used in various types of trials, ranging from nutritional, disease related, to toxicological and/or environmental, with the further possibility of comparing them at different time points (Zhang et al., 2012). This permits scientists to have a comprehensive view of animals' metabolism, which can provide insights into how to improve health, nutrition, production, reproduction, but also to study the final products delivered from livestock, such as meat, milk, as well as products therefrom (Singh et al., 2019).

1.4 Proteomics

Proteomics is defined as the study which aims at the characterization of the complete set of proteins present in a cell, organ, or organism in a specific time point (Wilkins et al., 1996). Proteomics approaches can be used to address a variety of scientific questions, such as (a) proteome profiling, (b) localization and identification of post-translation modifications, (c) comparison of proteins expression in different individuals, and (d) analyses of protein-protein interaction (Chandramouli and Qian, 2009). Proteomics thus aims to study the proteome, defined for the first time by (Wasinger et al., 1995) as the "total protein complement of a genome." The latter paper determined a transition between a one-protein-at-a-time analysis into a more global approach, in which proteins were starting to be studied in a large scale (Wilkins, 2009).

In the last years, proteomics (but in general all the "omics" approaches) has developed rapidly, also attributed to the improvement in technology and bioinformatics tools (Ceciliani et al., 2018a).

However, before giving an overview on the different types of analyses, another problem should be addressed: the concept of high- and low-abundance proteins. In mammals for example, the top ten most abundant proteins represent approximately 90% of the whole plasma proteome, while the other 10% is present in a very wide dynamic range (Liumbruno et al., 2010). Similarly, in cows' milk, the most abundant proteins are represented by the caseins which represent roughly 78% of the total protein concentration, with another 17% made up of whey proteins (Bendixen et al., 2011; Roncada et al., 2012). This can be problematic because some methods are able to detect and identify small numbers of proteins but fail in detecting and quantifying proteins which are less abundant (Millioni et al., 2011). For this reason, the targeted depletion of the most abundant proteins is often used to increase both the depth of proteome identification and to increase the sensitivity of targeted analyses for specific proteins (Tu et al., 2011).

Proteomics can be divided in a great variety of methods (Chandramouli and Qian, 2009), as shown in Table 4. Each of the methods has its own advantages and disadvantages.

Table 4. *Common proteomics technologies and their applications, strengths and limitations (adapted from Chandramouli and Qian, 2007)*

Technology	Applications	Strengths	Limitations
2DE	Protein separation, quantitative profiling	Relative quantification, sensitive to protein processing	Poor separation of acidic, basic, hydrophobic and low abundant proteins, large amount of sample required
DIGE	Protein separation, quantitative profiling	Relative quantification	Same as 2DE approaches, and difficulties in labelling proteins without lysine, requires special equipment

Protein array	Quantitates specific proteins for diagnostics (biomarkers or antibody detection) and discovery research	High-throughput, low sample consumption, high sensitivity	Small sample amounts: limited replication, targeted approach
Mass Spectrometry	Primary tool for protein identification and characterization	Functional analyses, data mining and knowledge discovery	No individual method to identify all the proteins, sensitivity is a problem
Bioinformatics	Analyses of qualitative and quantitative data	Functional analyses, data mining	No integrated pipeline for processing data

2DE: Two-dimensional electrophoresis
DIGE: Difference Gel Electrophoresis

1.4.1 Methods of analysis

In 1970, Laemmli (1970) firstly described a denaturing polyacrylamide gel system utilizing sodium dodecyl sulfate (SDS) to separate protein molecules based on size, which is the base of the One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE).

The first and earliest method for proteomics analyses is the two-dimensional electrophoresis (2DE), developed even prior to the definition of the term proteomics (Klose, 1975). This technique was the only one that could be routinely applied for parallel quantitative expression profiling of complex proteins (Görg et al., 2004). The 2DE requires the separation of complex proteins mixtures by molecular charge in the first dimension and by mass in the second dimension (Chandramouli and Qian, 2009). It is widely used for qualitative experiments, as its reproducibility is limited, and its inability to detect low abundance proteins is a drawback (Chandramouli and Qian, 2009).

A variation of the previous method is the Fluorescence 2D Difference Gel Electrophoresis (2D-DIGE), in which proteins are labelled with the fluorophores Cy2, Cy3 and Cy5 prior to the 2DE procedure (Ünlü et al., 1997). CyDyes are cyanine dyes that

carry an N-hydroxysuccinimidyl ester reactive group that binds the e-amino group of lysine residues in proteins. The limits of this method are the inability to label protein without lysine and the requirement for special equipment to pursue the analyses (Chandramouli and Qian, 2009). Nevertheless, the DIGE technique has higher sensitivity than the 2DE and is also more reproducible when comparing samples under similar electrophoretic conditions (Van Den Bergh and Arckens, 2005).

Protein arrays, also known as protein chips, are formed by immobilizing individually-purified proteins on a microscopic slide-based surface (Huang and Zhu, 2017). Depending on their applications, they can be classified into three categories: analytical, functional, and revere-phase protein arrays (Sutandy et al., 2013) – which will not be discussed as not part of the present thesis. Recently, protein arrays have seen a great development due to the development of new high-throughput methodologies, making these methods highly sensitive and capable of measuring hundreds of known proteins in different biological matrices, and enabling this methodology to be an important tool for quantitative proteomics studies, diagnostic discovery, and biomarkers development (Goshima et al., 2008; Chandramouli and Qian, 2009). The advantage of this method is the possibility to have a global overview of the proteome and to simultaneously screen for protein-protein interactions as well as post-translation modifications (Nijdam et al., 2009). However, the challenges of applying this method derive from the difficulty in creating a comprehensive expressions library, the management of the high-throughput large quantity of data, and the adaptation of DNA microarray to accommodate protein substrates (Phizicky et al., 2003).

Similarly to metabolomics analyses in which the aim is to identify the metabolites, the mass spectrometer is the primary tool when aiming at the identification of individual proteins (Chandramouli and Qian, 2009). Mass spectrometers consist of an ion source, a mass analyser, and an ion detection system. A complete analysis of proteins by mass spectrometers has three major steps: protein ionization and generation of gas-phase ions,

separation of ions according to their mass to charge ratio, and detection of ions (Mann et al., 2001). There are multiple methodologies to pursue a mass spectrometry study, as comprehensively reviewed by (Han et al., 2008), however, this is not in focus of this thesis and will thus not be addressed in detail.

The final step of a proteomics analyses is the bioinformatics part, in which the data collected from the analyses need to be evaluated. The major problem when analysing the results is that proteomics analyses can generate a huge amount of data, and there are different pipelines to evaluate these data, and not all of them are capable of generating real biological insights (Chandramouli and Qian, 2009).

As an example, the classical workflow for processing proteomics (Figure 5) data sets would include:

- Sample collection and protein extraction;
- Data acquisition (MS-based proteomics);
- Analysis: data dependent or independent acquisition;
- Data processing (single-shot-proteins, number of peptide per protein, features analysis);
- Statistical analysis (supervised or unsupervised analysis)
- Biological interpretation (e.g., pathway analysis)

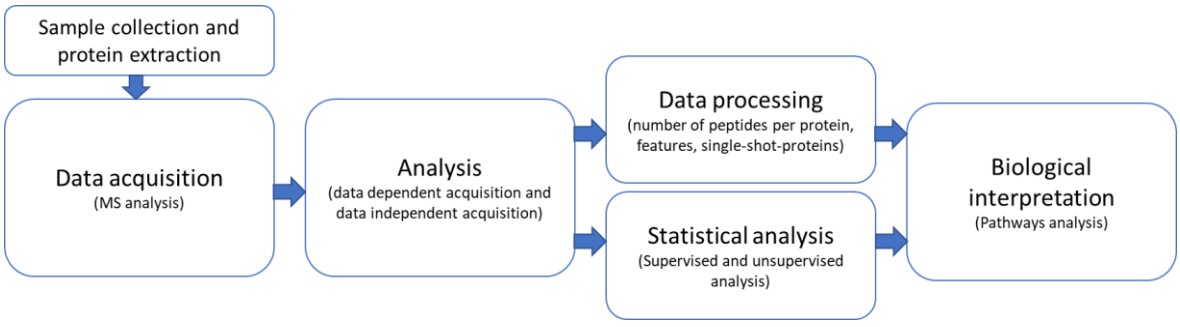


Figure 5. *Example of a typical workflows for untargeted proteomics analyses*

1.4.2 Applications of proteomics in animal science

As previously addressed in the metabolomics section, “omics” approaches are now used to obtain a more complete picture of system-wide metabolism and biology. Proteomics can thus be used to obtain a snapshot on the condition of cells, tissues, or body fluids in a specific time point (Ceciliani et al., 2018a). For this reason, proteomics analyses in animal sciences have recently increased, even though they still represent a minority in the proteomics world (Bilić et al., 2018). Proteomics is currently used in a variety of studies, aiming at researching energy metabolism regulation, parasite tolerance, disease response, or feed efficiency (Soares et al., 2012; Bilić et al., 2018). As proteins are involved in many metabolic processes of an individual, studying the proteome is becoming fundamental in order to fully understand the biochemical and physiological aspects of farm animal biology and its relation to health and productivity (Soares et al., 2012).

In future research, as it was outlined in the work of Chait (2011), there is still the need to improve sample handling and sensitivity of detection to allow the analysis of smaller samples, and at the same time develop even further methodologies for measuring low abundance proteins. Moreover, as reported by Almeida et al. (2021), proteomics analyses may play a pivotal role in animal science research focussed on environment and sustainability, being the perfect tool to complement the traditional research areas (nutrition, genetics, physiology, and others).

Chapter 2 – Objectives

The body condition of dairy cows is known to be related with animal health and performance. Cows entering lactation in an over-conditioned status are more prone to lose body reserves and thus susceptible to metabolic distress which may result in the development of various health disturbances and diseases. So far, most of the studies investigating the relation between body fat and metabolic adaptations during the transition period have classified cows by using arbitrarily defined thresholds for BCS or BFT. However, this is not necessarily the most appropriate and definitive way of grouping and is also inefficient when studying large data sets.

To further study the variation during the transition period of dairy cows, our objectives were:

(1) To select by means of machine learning (to reduce at the minimum a human selection which could results biased and unprecise) a representative number of dairy cows, from which samples comprising the most informational value based on performance, health and BFT data could be derived,

(2) To characterize the inter-individual variation in the relationship between pre-calving BFT and subsequent BFT loss during early lactation in a large dairy herd, in terms of milk production, health condition, and circulating concentration of metabolites (NEFA, BHB), metabolic hormones (leptin and adiponectin), and an inflammatory marker (haptoglobin) both *ante* and *post partum*; all these variables are important during the transition period as they can be indicator of the metabolic status of dairy cows around calving, but also,

(3) To apply metabolomics and proteomics analyses to have a complete overview on the metabolome and proteome. Specifically, subgroups of animals (selected as the most representative ones) were identified and tested to determine whether cows differing in body condition *ap* and/or with a different mobilization around calving would differ when

analysing the serum proteome and metabolome around 30 DIM, respectively. Moreover, a metabolomics analysis was carried out to study whether the effect of a Methionine supplementation when considering animals with different body condition and/or with a different clinical status (i.e., affected or not by endometritis).

We hypothesized that over-conditioned and/or high mobilizing dairy cows around calving will be different in their subsequent metabolic changes from normal and under-conditioned cows when looking at classical variables, and in addition we aimed to evaluate the differences of dairy cows' metabolome and proteome during their *pp* depending on their BC loss, diet, and BFT *ap*, respectively.

Chapter 3 – Identification and characterization of dairy cows with different ante-partum backfat thickness (BFT) in relation to BFT loss post-partum: A cluster analytic approach

Ruben Riosa contributions: data analyses, draft of the manuscript

Published Article: R. Riosa, M. H. Ghaffari, H. M. Hammon, D. Süss, M. Hoelker, M. Drillich, C. Parys, J. Guyader, H. Sauerwein, and M. Iwersen. **Identification and characterization of dairy cows with different ante-partum backfat thickness (BFT) in relation to BFT loss post-partum: A cluster analytic approach**

Co-author contributions: MHG, supported data analyses, and drafting and revision of the manuscript; HMH, analyses of BHB and NEFA, revision of the manuscript; DS, sample collection; MD, study design, revision of the manuscript; CP, study design; JG, revision of the manuscript; HS, study design, revision of the manuscript; MI, study design, sample and data collection, and revision of the manuscript.

3.1 Abstract

The objectives of this study were (1) to characterize the interindividual variation in the relationship between *ap* backfat thickness (BFT) and subsequent BFT loss during early lactation in a large dairy herd using cluster analysis, (2) to compare the serum concentrations of metabolites (non-esterified fatty acids, ß-hydroxybutyrate), metabolic hormones (leptin and adiponectin), and an inflammatory marker (haptoglobin) among the respective clusters, and (3) to compare lactation performance and uterine health status in the different clusters. An additional objective was (4) to investigate differences in the serum variables mentioned above and in milk yield of over-conditioned (OC) cows that differed in the extent of BFT loss. Using data from a large study of 1,709 multiparous Holstein cows, we first selected those animals from which serum samples and BFT results (mm) were available at d 25 (± 10) ap and d 30 (± 3 d) *pp*. The remaining 713 cows (parity of 2 to 7) were then subjected to cluster analysis: different approaches based on BFT of the cows were performed. K-means (unsupervised machine learning algorithm) clustering based on BFT-ap alone identified five clusters: lean (5-8 mm BFT, n = 50), normal (9 - 12 mm, n = 206), slightly fat (SF; 13 - 16 mm, n = 203), just fat (JF; 16 - 22 mm, n = 193), and very fat (VF; 23 - 43 mm, n = 61). Clustering by difference between BFT-ap and BFT pp (ΔBFT) also revealed five clusters: extreme loss (17 - 23 mm ΔBFT, n = 16), moderate loss (9 - 15 mm, n = 119), little loss (4 - 8 mm, n = 326), no loss (0 - 3 mm, n = 203), and gain (-8 to -1 mm, n = 51). Based on the blood variables measured, our results confirm that cows with greater BFT losses had higher lipolysis and ketogenesis than cows with less BFT loss. The serum variables of cows that gained BFT did not differ from normal cows. Milk yield was affected by the BFT-ap cluster, but not by the ΔBFT cluster. Cows categorized as VF had lesser milk yield than other clusters. We further compared the OC cows that had little or no BFT loss (i.e., 2% of VF, 12% of JF, and 31% of SF, OC-no loss, n = 85) with the OC cows that lost BFT (OC-loss, n = 135). Both NEFA and BHB pp

concentrations and milk yield were greater in OC-loss cows compared with the OC-no-loss cows. The serum concentration of leptin ap was greater in OC-loss than in the OC-no-loss cows. Overall, OC cows lost more BFT than normal or lean cows. However, those OC cows with a smaller loss of BFT produced less milk than OC cows with greater losses.

Key words: dairy cows, cluster analysis, transition period, body condition loss

3.2 Introduction

Dairy cows face physiological challenges during the transition from late pregnancy to lactation (Drackley, 1999). In early lactation, dairy cows are typically in a state of negative energy balance (NEB) because feed intake is insufficient to meet the increased nutrient requirements for milk synthesis (Drackley, 1999). Although metabolizing body fat for milk production is very important, it is known that cows with a high body condition score (BCS) around calving are at greater risk of developing metabolic disorders, mainly ketosis, and impaired fertility (Bernabucci et al., 2005; Roche et al., 2009; Rathbun et al., 2017). In addition to the BCS around calving, the magnitude of BCS loss during the transition from pregnancy to lactation may be even more important for metabolic health than BCS alone (Rathbun et al., 2017). A loss of one unit of BCS on a 5-point-scale double the risk of ketosis (Duffield et al., 1998; Rathbun et al., 2017). Excessive loss of body energy reserves has also been associated with impaired immune function, increasing the risk of infectious diseases such as endometritis (Esposito et al., 2014). For milk yield, reports in the literature are conflicting, and as comprehensively reviewed by (Roche et al., 2009), there is no linear relationship between BCS at calving and milk yield. Body condition around calving is considered as the main determinant of BCS loss, i.e., the greater BCS is at calving, the more body reserves are mobilized pp, increasing the risk of metabolic disorders (Roche et al., 2007b). Therefore, it is generally recommended to avoid over-conditioning before calving and to protect against excessive losses during the transition period.

In addition to BCS, body fat reserves can also be assessed by ultrasonic measurements of the backfat thickness (BFT; Schröder and Staufenbiel, 2006). The accumulation of body fat reserves and the extent to which these reserves are mobilized varies among individuals. The reasons for this variation are largely unknown. Using plasma proteomics, it has been shown that over-conditioning around calving is associated

with changes in signalling pathways related to the acute inflammatory response and regulation of complement and coagulation cascades (Ghaffari et al., 2020b). Furthermore, using targeted metabolomics, divergent metabotypes have been identified even within well-characterized phenotypes with high BCS and BFT at calving and severe subsequent BCS and BFT loss (Ghaffari et al., 2020a). Thus, individual cows with comparable body fat portions appear to differ in how they manage their fat reserves during the transition period, even when fed and managed in the same manner.

With this background, we used a cluster analytic approach on 713 animals to identify clusters of cows based on either *ap* BFT (BFT-ap) or the BFT change during the transition period (BFT loss, ΔBFT), with thresholds that were not arbitrary but were derived solely from the data set used. We used data and samples from a previously described feeding trial (Süss et al., 2019). The objectives of the present study were (1) to characterize the interindividual variation in the relationship between ap BFT and subsequent BFT loss during early lactation in a large dairy herd, (2) to estimate into which ΔBFT cluster cows from the different BFT-ap clusters would develop, and (3) to compare milk yield, uterine health (mainly prevalence of endometritis and hyperketonaemia), and metabolic and inflammatory status (serum variables: non-esterified fatty acids, ß-hydroxybutyrate, leptin, adiponectin, and the acute phase protein haptoglobin) among the different clusters. The descriptive comparisons of cows from a large dairy herd with divergent BFT ap status extend the knowledge about the biological variation in body condition ap and their after-effects on the mobilization of body reserves. The information about this variation within the same herd kept under the same management conditions on a commercial dairy farm is relevant for researchers and dairy farmers for finding strategies of either eliminating less efficient animals or aiming at “personalized” feeding and management systems.

3.3 Material and Methods

This study was approved by the Slovakian Regional Veterinary Food Administration, and by the institutional ethics committee of the University of Veterinary Medicine, Vienna, Austria (ETK-09/02/2016). The animal experiment was conducted from March 2016 to November 2017 in a commercial dairy farm in Slovakia, where about 2,400 Holstein-Friesian cows are kept.

3.3.1 Animals, diets, records, and samplings

The current study extends the study by (Süss et al., 2019) in which the effect of dietary supplementation with rumen-protected methionine (Met) on reproductive performance was tested in 1,709 multiparous Holstein cows. All animals received the same basal diet as a total mixed ration (TMR) during pregnancy and lactation as reported previously (Süss et al., 2019). Cows were offered the TMR twice daily and adjusted to achieve refusal rates of 5 to 10%. After calving and leaving the fresh group (between d 8 and 40 *pp*), the cows were assigned to either a basal diet (control) or the basal diet plus approximately 27 g/d of rumen-protected Met (Mepron®, Evonik Operations GmbH, Hanau, Germany) which was added to the mineral and vitamin premix and fed in the TMR until the end of the observation period (70 d pp).

As represented on Figure 6, on d 25 ± 10 ap and d 31 ± 3 pp (means ± SD), BFT was assessed by ultrasound (Easi-Scan, IMV imaging, Bellshill, Scotland, UK; 7.5 MHz linear probe with wrist display) as previously described (Schröder and Staufenbiel, 2006). Cows were evaluated for metritis at d 5 pp by assessment of vaginal discharge and body temperature, and for endometritis at d 31 ± 3 pp by vaginal examination and uterine cytology (Süss et al., 2019). Endometritis was diagnosed using a modified vaginal discharge score (Williams et al., 2005) to classify vaginal mucus as (E0) clear mucus; (E1) ≤ 50% non-white or white mucopurulent material; and (E2) ≥ 50% off-white or white

mucopurulent material. Furthermore, uterine cytology samples were collected using the Cytobrush method, as described by (Kasimanickam et al., 2004). These samples were then prepared by rolling the brush onto a clean glass slide. On the farm, the slides were immediately fixed and stained (LT-SYS, Labor und Technik, Berlin, Germany) and evaluated under the microscope (x 400 magnification) by counting a total of 300 cells to determine the percentage of polymorphonuclear neutrophils (PMN; (Melcher et al., 2014). The cutoff point for diagnosing subclinical endometritis was 5% PMN (Madoz et al., 2013). Based on the vaginal examination and uterine cytology, cows were classified as healthy (vaginal discharge score = E0, proportion of PMN < 5%) or affected by endometritis: subclinical endometritis (SE; vaginal discharge score = E0, proportion of PMN ≥ 5%) or clinical endometritis (CE; vaginal discharge ≥ E1).

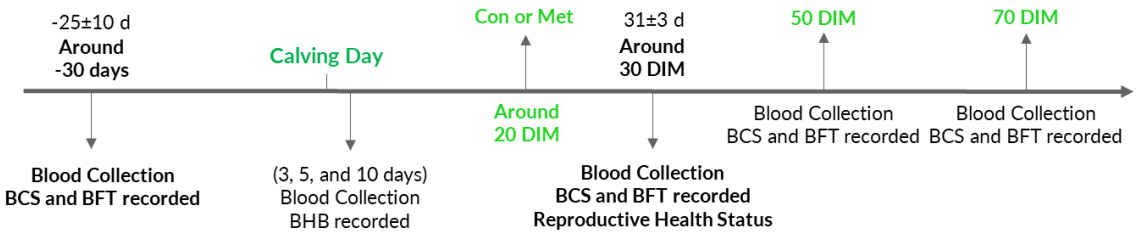


Figure 6. *Representation of the sampling and data collection from the animal study*

Blood samples were collected from a coccygeal vessel using vacuum tubes coated with a clot activator for serum collection (Süss et al., 2019), on the same day as the BFT measurements. In the samples, serum ß-hydroxybutyrate (BHB) was used to classify cows as normal or hyperketonaemia using a BHB threshold of 1.2 mM (Süss et al., 2019). Two metabolic hormones (leptin and adiponectin), an acute-phase protein (haptoglobin), and two metabolites [BHB and non-esterified fatty acids (NEFA)] were measured in serum as described below. Cows were milked twice daily in a rotating milking parlor and milk production was recorded daily throughout the study. Milk production (kg/d) is expressed as a weekly average. Only cows from which all records and samples were available and

which were assigned to the control group or Met treatment no later than d 20 pp were included in the present study, resulting in a total of 713 cows.

3.3.2 Clustering of the Cows according to BFT and to BFT Loss

The k-means (unsupervised machine learning algorithm) analysis procedure of the SAS package (PROC FASTCLUS; SAS Institute 253 Inc, Cary, NC) was used with the mean, median, sum, and standard deviation of BFT at 25 ± 10 d ap (BFT-ap) or BFT loss (delta-BFT; ΔBFT = BFT at d 25 ap minus BFT at d 31 pp) as key features for each cluster procedure. The number of clusters was determined by tests based on the separation between clusters and homogeneity within clusters. To validate the clusters, we estimated the optimal number of clusters for k-means clustering using the direct method in R (V 4.0.3; R Core Team, 2019). This method aims to optimize a criterion, the sum of squares within clusters or the average silhouette, using the method of elbows and average silhouette (Rousseeuw, 1987). Statistical power was defined as the probability of the cluster analysis to reject the null hypothesis (no clustering found), which was defined by an average silhouette width above 0.5 (Kaufman and Rousseeuw, 1990).

3.3.3 Laboratory Analysis of all 713 cows (ap and pp samples)

The concentrations of BHB and NEFA were measured spectrophotometrically (HORIBA ABX SAS, Montpellier, France) at the Research Institute for Farm Animal Biology (FBN) in Dummerstorf (Germany) using the following kits: BHB (#RB1008, Randox Laboratories Limited, County Antrim, United Kingdom) and NEFA (#434-91795, Wako Chemicals GmbH, Neuss, Germany). The two kits are formed by semi-quantitative dipstick which can detect acetone and acetoacetate for the BHB, and a quantitative test for the detection of NEFA. Depending on the quantity detected, the different colours will be

detected by the spectrophotometer, resulting with the concentration of the two metabolites in the serum. Adiponectin, leptin, and haptoglobin were measured using ELISA methods developed in-house (Sauerwein et al., 2004; Hiss et al., 2009; Mielenz et al., 2013). For the adiponectin ELISA, the intra- and inter-assay coefficients of variation (CV) were 9.8 and 13.4%, respectively. The intra- and inter-assay CV were 9.2 and 13.4% for leptin and 8.9 and 11.6% for haptoglobin, respectively.

3.3.4 Statistical analyses

A repeated-measures model was fitted to the data (both blood and milk) using a linear mixed-effects model with SPSS software (IBM® SPSS® Statistics 24.0, Armonk, NY, USA), using treatment (BFT-ap clusters, ΔBFT clusters, or BFT-loss), time (d, wk), time × treatment interaction, MET supplementation (control or MET), uterine health status (healthy, subclinical, or clinical endometritis), parity as fixed effects, and cow as a random effect. In the preliminary data analysis, 2-way interactions (treatment × uterine health status; treatment × Met supplementation; time × Met supplementation; time × uterine health status) were included in the model as a fixed effect, but the result of this preliminary statistical analysis showed no significant effect of the 2-way interactions on the tested variables in this study. Therefore, these 2-way interactions were disregarded in the final statistical analysis of the data. All data were tested for normal distribution using the UNIVARIATE procedure of SAS (Shapiro-Wilk test), and variables that were not normally distributed were $\log_{10}$-transformed to meet the assumptions of normality and homoscedasticity of the residuals. Data are presented as means ± SEM, and Tukey-Kramer adjustment was applied to account for multiple comparisons. Significance was declared at $P \leq 0.05$ and a trend at $0.05 < P \leq 0.10$.

3.4 Results

3.4.1 Characteristics of the clusters obtained for BFT-ap

Five clusters were determined based on BFT-ap: lean (LEN, 5 to 8 mm BFT), normal (NOR, 9 to 12 mm BFT), slightly fat (SF, 13 to 16 mm BFT), just fat (JF, 16 to 22 mm BFT), and very fat (VF, 23 to 43 mm BFT). Figure 7A shows a histogram indicating the frequency (number of animals) in the obtained clusters. The Gaussian distribution showed anomalous skewness and kurtosis coefficients, and the Shapiro-Wilk test for the obtained clusters was significant, indicating that the data were not normally distributed (Figure 7A). There were fewer cows in the LEN (n = 50, 7%) and VF (n = 61, 8.6%) clusters, with the majority observed in NOR (n = 206, 28.9%), SF (n = 203, 28.5%), and JF (n = 193, 27.1%).

For the different BFT-ap clusters, milk yield and serum parameters are presented in Figures 7B and 7C, respectively. Milk yield was affected by BFT-ap cluster, time, parity, and uterine health ($P < 0.01$). Cows categorized as VF had lesser milk yield than other clusters. Milk yield increased with time ($P < 0.01$) and reached a plateau around 7-8 wk of lactation. The serum concentration of BHB was affected by BFT-ap cluster ($P < 0.01$), time ($P < 0.01$), cluster × time interaction (trend, $P = 0.06$), and Met ($P = 0.02$). The concentration of NEFA was affected by BFT-ap cluster ($P < 0.01$), time ($P < 0.01$), cluster × time interaction (trend, $P = 0.08$), parity ($P < 0.01$). The concentrations of NEFA and BHB after calving were greater in VF and JF compared with the other clusters. The leptin concentrations were affected by BFT-ap cluster and time ($P < 0.01$; Figure 7C). Compared with the other treatment groups, the VF group had the highest concentration of ap-leptin, whereas the LEN group had the lowest. The adiponectin concentrations in serum were affected only by time and decreased after calving. The concentrations of Hp in serum were affected by the main effects or by the cluster × time interaction. As expected, circulating

concentrations of NEFA and BHB increased, whereas leptin and adiponectin decreased from ap to pp.

In addition, no differences were found among clusters when other factors, including Met treatment, health status, and parity were considered (Supplemental Table S1). The incidence of endometritis was assessed at d 30 pp, but we did not observe any differences among the 5 different BFT-ap clusters, neither in uterine health status nor in calving and conception intervals (Supplemental Table S2).

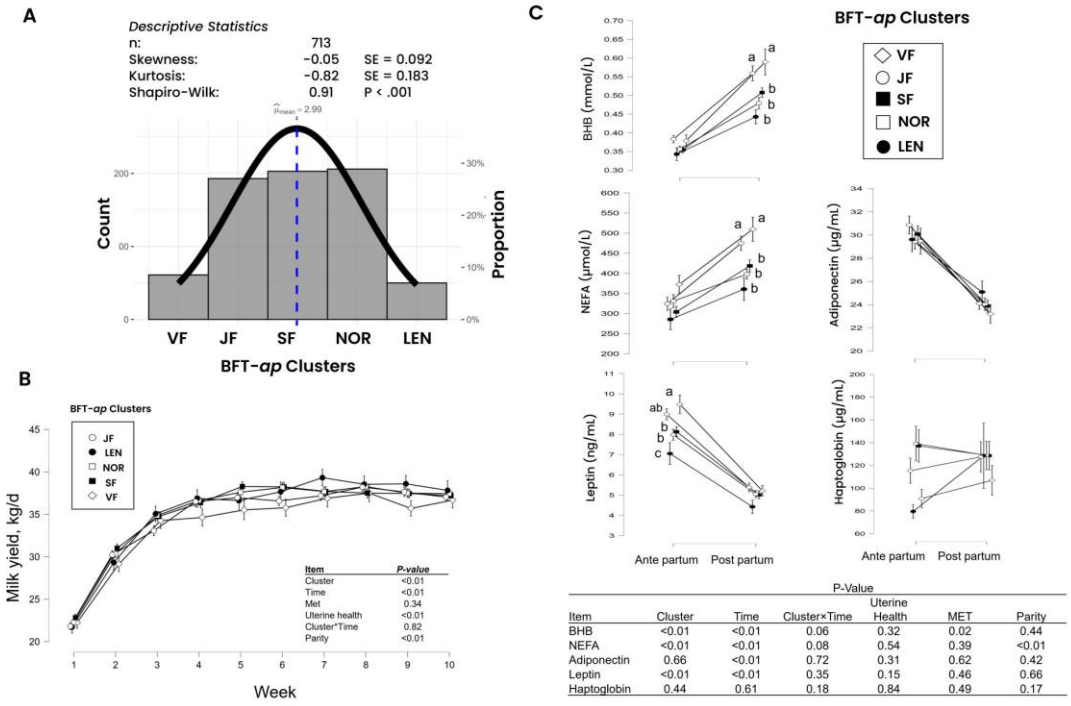


Figure 7. *Characteristics of the clusters based on backfat thickness ap (BFT-ap): LEN: lean, BFT 5 to 8 mm; NOR: normal, BFT 9 to 12 mm; SF: slightly fat, BFT 13 to 16 mm; JF: just fat, 16 to 22 mm; VF: very fat, 23 to 43 mm. (**A**) histogram of the distribution, and (**B**) milk yield (means ± SEM), (**C**) serum concentration (means ± SEM) of various metabolites (NEFA and BHB), metabolic hormones (leptin and adiponectin) and of the acute phase protein (haptoglobin) in samples collected ap (30 d before calving) and pp (30 d after calving). Uterine health = cows were classified as healthy or affected by endometritis based on vaginal examination and uterine cytology. Vaginal discharge score = clear mucus proportion of PMN < 5% or affected by endometritis: subclinical endometritis (vaginal discharge score = ≤ 50% off-white or white, proportion of PMN ≥ 5%) or clinical endometritis (≥ 50% off-white or white mucopurulent material). PMN = proportion of polymorphonuclear neutrophils.*

3.4.2 Characteristics of the clusters obtained for ΔBFT

When the k-means clustering was based only on ΔBFT, 5 clusters were obtained, as shown in Figure 8A. Evaluation of the Gaussian distribution showed anomalous skewness and kurtosis coefficients, and the p-value of the Shapiro-Wilk test for the clusters obtained for ΔBFT was less than 0.01, indicating that the data were not normally distributed (Figure 8A). A smaller number of cows were observed in the extreme loss (EL, n = 16, 2.2%) and gain (GN, n = 51, 7.2%) clusters, whereas the majority of cows were found in the small loss (SL, n = 326, 45.7%), moderate loss (ML, n = 119, 16.7%), and no loss (NL, n = 201, 28.2%) clusters.

Figure 8 shows milk yield (Figure 8B) and serum variables (Figure 8C) associated with the ΔBFT clusters. Milk yield was affected by time, parity, and uterine health ($P < 0.01$), but not by the ΔBFT cluster, Met, or the cluster × time interaction. The BHB concentration in serum was affected by ΔBFT cluster ($P < 0.01$), time ($P < 0.01$), cluster × time interaction ($P = 0.02$), and MET ($P = 0.02$) but not by parity and uterine health. The NEFA concentrations were influenced by ΔBFT cluster ($P < 0.01$), time ($P < 0.01$), cluster × time interaction ($P = 0.01$), and parity ($P < 0.01$), but not by Met and uterine health. The NEFA and BHB concentrations pp were greater in EL and ML than in the other clusters. Serum leptin concentrations were affected by time ($P < 0.01$) but not by ΔBFT cluster, cluster × time interaction, parity, Met, and uterine health. Serum adiponectin concentrations were affected only by the ΔBFT cluster and cluster × time interaction and were lower in EL cows compared with other clusters ap. Serum Hp concentrations were affected by the main effects, and by the cluster × time interaction.

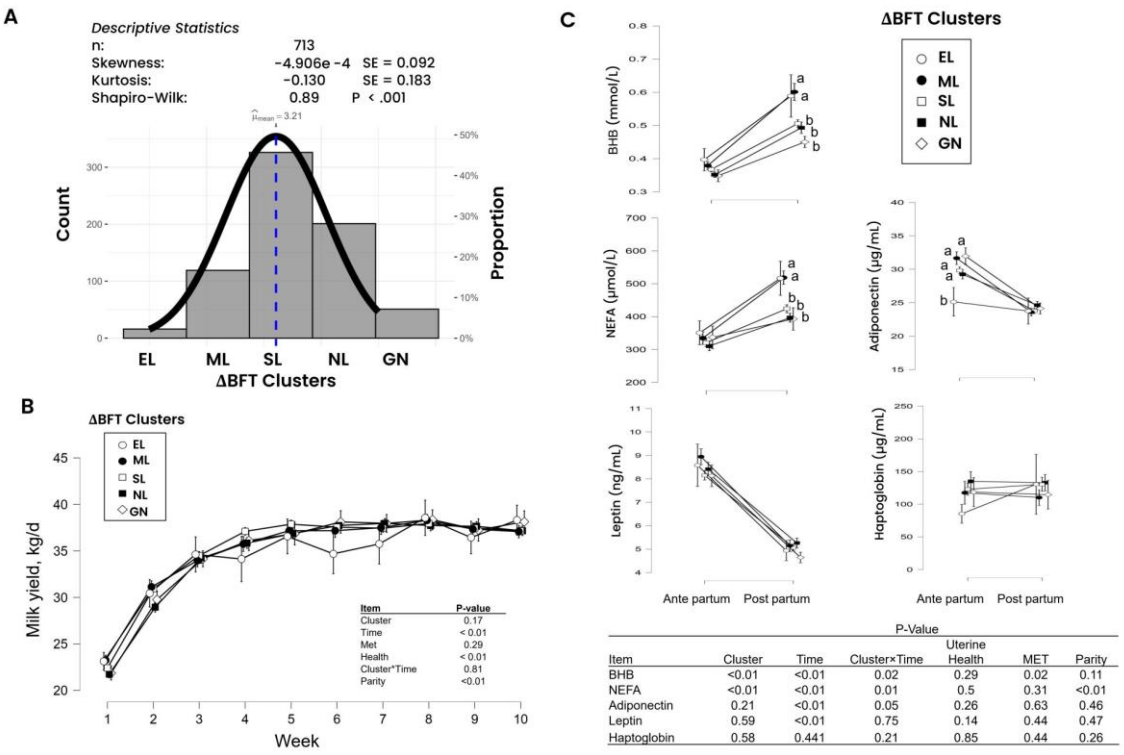


Figure 8. *Characteristics of the clusters based on backfat thickness loss from ap to pp (ΔBFT): EL: extreme loss (ΔBFT from 17 to 23 mm); ML: moderate loss (9 to 15 mm); SL: small loss (4 to 8 mm); NL: no loss (0 to 3 mm); GN: gain (-8 to -1 mm). (**A**) histogram of the distribution, and (**B**) milk yield (means ± SEM), (**C**) serum concentration (means ± SEM) of various metabolites (NEFA and BHB), metabolic hormones (leptin and adiponectin) and of the acute phase protein (haptoglobin) in samples collected ap (30 d before calving) and pp (30 d after calving). Uterine health = cows were classified as healthy or affected by endometritis based on vaginal examination and uterine cytology. Vaginal discharge score = clear mucus proportion of PMN < 5% or affected by endometritis: subclinical endometritis (vaginal discharge score = ≤ 50% off-white or white, proportion of PMN ≥ 5%) or clinical endometritis (≥ 50% off-white or white mucopurulent material). PMN = proportion of polymorphonuclear neutrophils.*

3.4.2 Apportionment of the cows in the BFT-ap clusters to the ΔBFT clusters

To illustrate the presumed dependence of BFT loss on BFT-ap, we used a Sankey plot diagram (Figure 9) for showing the proportion of dairy cows from the different BFT-ap clusters that were classified into the different ΔBFT clusters. We observed that 98% of VF cows, 88% of JF cows, and 69% of SF cows were assigned to the ΔBFT clusters with greater losses compared with NOR or LEN cows. We further used the Sankey plot results to compare the over-conditioned (OC) cows that had little or no BFT loss (i.e., the 2% of VF, the 12% of JF, and the 31% of SF; n = 85) with the OC cows that lost BFT (n = 135).

Figure 9. *Sankey plot for representing the apportionment of animals in five BFT-ap clusters [LEN: lean, BFT 5 to 8 mm; NOR: normal, BFT 9 to 12 mm; SF: slightly fat, BFT 13 to 16 mm; JF: just fat, 6 to 22 mm; VF: very fat, 23 to 43 mm.] to the five ΔBFT clusters [EL: extreme loss (ΔBFT from 17 to 23 mm); ML: moderate loss (9 to 15 mm); SL: small loss (4 to 8 mm); NL: no loss (0 to 3 mm); GN: gain (-8 to -1 mm)].*

The resulting groups, designated as OC with little or no BFT loss (OC-no loss) and OC with severe to normal loss (OC-loss), were then compared for milk yield (Figure 10A) and their serum concentrations of the assessed variables (Figure 10B). Milk yield was affected by the OC-ΔBFT (OC with different BFT losses) groups, time, parity, and uterine health ($P < 0.01$), but not by Met supplementation, or the group × time interaction. Milk yield was greater in the OC-loss cows than in OC-no-loss cows ($P < 0.01$). The BHB concentration in serum was affected by the OC-ΔBFT group ($P < 0.01$), time ($P < 0.01$), group × time interaction (trend, $P = 0.09$), and Met (trend, $P = 0.07$) but not by parity and uterine health. The NEFA concentration was affected by the OC-ΔBFT group ($P < 0.01$), time ($P < 0.01$), group × time interaction (trend, $P = 0.10$), parity (trend, $P = 0.07$), but not by Met and uterine health. The pp concentrations of NEFA and BHB were greater in OC-loss than in the OC-no-loss cows. Serum adiponectin was affected only by time ($P < 0.01$) and leptin was affected by the OC-ΔBFT group, time, and uterine health ($P < 0.01$), but not

by the group × time interaction, parity, and Met. The ap concentration of leptin was greater in OC-loss than in the OC-no loss cows, whereas none of the fixed effects tested was significant for Hp.

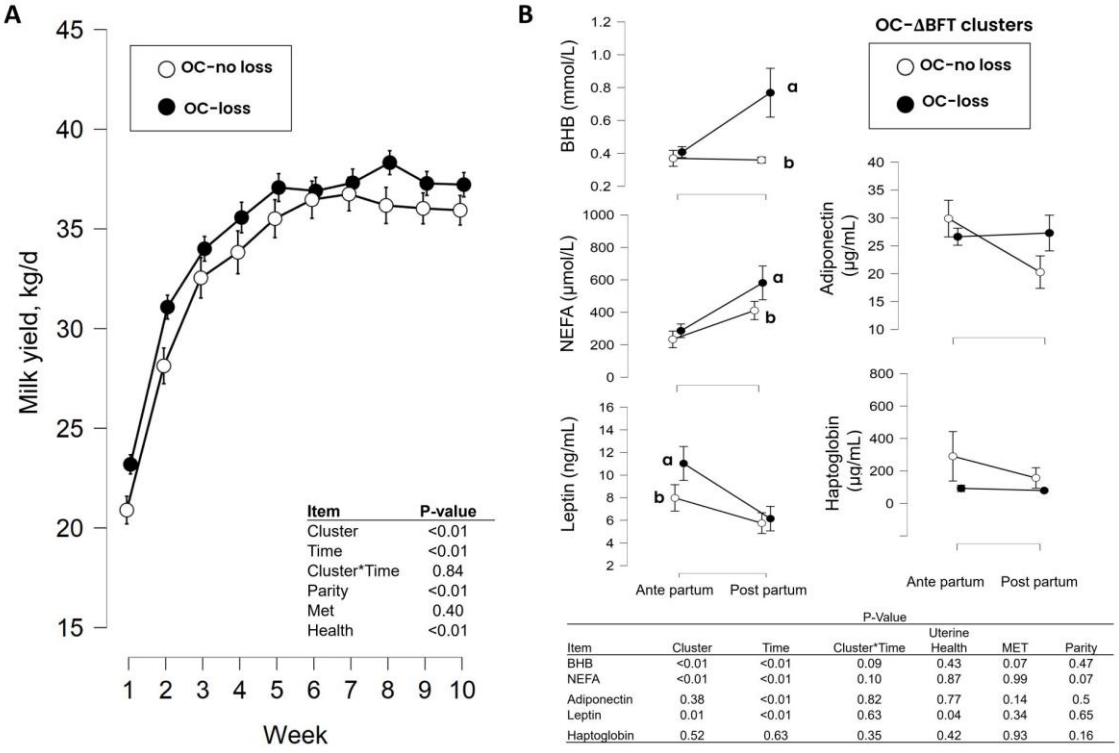


Figure 10. *(**A**) Milk yield (means ± SEM), (**B**) serum concentration (means ± SEM) of various metabolites (NEFA and BHB), metabolic hormones (leptin and adiponectin) and of the acute phase protein haptoglobin in samples collected ap (30 d before calving) and pp (30 d after calving) in overconditioned cows with different backfat thickness (BFT) loss. OC-no-loss = overconditioned cows with little or no BFT loss, OC-loss = overconditioned cows with severe to normal loss. Uterine health = cows were classified as healthy or affected by endometritis based on vaginal examination and uterine cytology. Vaginal discharge score = clear mucus proportion of PMN < 5% or affected by endometritis: subclinical endometritis (vaginal discharge score = ≤ 50% off-white or white, proportion of PMN ≥ 5%) or clinical endometritis (≥ 50% off-white or white mucopurulent material). PMN = proportion of polymorphonuclear neutrophils.*

Supplemental Figure S1 shows the frequency of cows supplemented with or without rumen-protected Met in different clusters. A description of the distribution of endometritis, and hyperketonaemia, as described by Süss et al. (2019), after our clustering based on BFT-ap and ΔBFT is shown in the Supplemental Tables S1 and S3, respectively. The frequency of cows with different parities in the different clusters is shown in Supplemental Figure S2. The calving to conception intervals (CI) calculated for the BFT-ap clusters for both the lactation cycles preceding and following the current study are presented in Supplemental Table S2. In summary, the portion of cows receiving Met supplementation ranged from 37.5 to 58.8% in all clusters (mean ± SD: 49.1 ± 7.9%), and

there was also no apparent shift in the proportion of animals with diagnoses among clusters.

3.5 Discussion

This study was conducted to characterize individual variation in the relationship between ap BFT and subsequent loss of BFT during early lactation in large dairy herds and to estimate which ΔBFT clusters cows from the different ap BFT clusters would develop. In addition, we examined milk yield, uterine health (focusing on endometritis and hyperketonaemia), and metabolic and inflammatory status among the different clusters. The extent of body fat mobilization during this transition varies between cows and usually shows positive correlations with the ap body condition of the animal (Weber et al., 2013). In the current study, there were no differences between the ΔBFT clusters for milk yield but there was a significant difference in milk yield of the clusters obtained for BFT-ap. In our study, VF cows produced less milk compared to the other clusters. This might have been due to lower breeding values for milk yield in these cows. Moreover, less milk yield in the preceding lactation might, in turn, result in a more positive energy balance thus allowing the animals to accumulate more body fat. Unfortunately, neither breeding values nor milk yields and feed intake of the preceding lactation were available in this study. There is no conclusive evidence on how BCS may impact milk yield in dairy cows. Several studies examining the effect of BCS on milk yield found no significant association, whereas others found an association between BCS and milk production. For example, in an analysis of 2,463 lactation records of pasture-based dairy cows by (Roche et al., 2007c), milk yield was found to be nonlinearly related to BCS at calving and the highest milk yield was found at a calving BCS of 3.50. A positive relationship between BCS loss (between calving and nadir) and milk yield was observed by (Roche et al., 2007c). In contrast, (Gobikrushanth et al., 2019) reported no association between pp BCS change categories

and cumulative milk yield in early lactation up to 25 and 90 days in milk, but peak and 305-d milk yield were greater in cows with extreme losses (loss ≥ 0.75 BCS units) compared to cows gaining BCS (≥ 0.25 units). Other studies found no association of BCS at calving or BCS change during early lactation with daily milk yield or daily energy-corrected milk yield in dairy cows [Pires et al., 2013 (up to 7 wk pp); Carvalho et al., 2014 (up to 3 wk pp)].

As shown in the Sankey plot, we observed that most OC cows (98% of VF and 88% of JF cows) were assigned to the ΔBFT clusters with greater losses compared with NOR or LEN cows. Furthermore, in this study, OC cows with little or no BFT loss (OC-no loss) produced less milk than OC cows with severe to normal losses (OC-loss). Likely, OC cows that do not lose their BCS during early lactation do not have the genetic merit to produce more milk. Smith and McNamara (1990) reported that cows with higher genetic merit for milk production experience greater lipolysis and mobilization of body reserves during early lactation, resulting in greater BCS loss and NEB.

It is well documented that with increased lipolysis and release of fatty acids from adipose tissue into the bloodstream, the hepatic capacity for FA oxidation is exceeded, leading to increased ketogenesis and also fatty liver (McFadden, 2020; Ghaffari et al., 2021). In the current study, we observed greater blood concentrations of NEFA and BHB in VF and JF as well as EL and ML cows compared with the other clusters. Moreover, also the cows with severe to normal loss (OC-loss) had greater NEFA and BHB concentrations than OC-no-loss cows, thus indicating a more negative energy balance and a greater allocation of energy to the mammary gland to produce more milk. However, feed intake could not be assessed in this study and thus we cannot substantiate the potential relationship of BFT loss with feed intake and energy balance.

The circulating concentrations of BHB and NEFA have been reported to reach greater levels in multiparous than in primiparous cows (Meikle et al., 2004). In multiparous

cows, (Nowroozi-Asl et al., 2011) also reported that greater incidences of subclinical ketosis in cows with more lactations. In our study, we observed that the BHB and NEFA values increased from parity 2 to 5 but were lesser in parity 6 and 7 compared to parity 5. However, the lesser values in parity 6 and 7 could be due to the relatively small sample size in our analysis.

In the current study, serum leptin concentrations decreased after calving in all clusters, as shown in previous studies (Reist et al., 2003; Kokkonen et al., 2005; Schuh et al., 2019), which is associated with the onset of NEB and a decrease in feed intake (Chilliard et al., 2005). The high leptin concentrations in VF cows are consistent with the notion that adiposity is an important determinant of leptinaemia (Chilliard et al., 2005). Similarly, in a previous study, ap plasma leptin concentration was highest in the high BCS group and did not differ between the moderate and low BCS groups (Pires et al., 2013). However, no BFT-ap effects were observed on leptin concentrations after calving, as previously observed (Holtenius et al., 2003; Pires et al., 2013).

In the current study, serum adiponectin concentration decreased after calving. The decrease in blood adiponectin concentration and lower mRNA abundance in tissues after calving may be due to increased lipolysis (Singh et al., 2014). In the current study, serum adiponectin concentration was affected only by the ΔBFT cluster and was lower in EL cows compared to the other clusters. Our results suggest that BFT is more closely related to leptin than adiponectin in dairy cows during the ap period, but this relationship might disappear as parturition approaches because adiponectin and leptin concentrations decrease after calving. Therefore, during early lactation, BFT loss associated with NEB may become more important than the degree of ap body fat in determining adiponectin concentration (Giesy et al., 2012; Singh et al., 2014; De Koster et al., 2017). This could explain why the degree of BFT loss was not associated with ap leptin concentration in our experiment.

Besides being an acute-phase protein (APP), Hp has also been identified as an adipokine in cattle (Saremi et al., 2012). However, given the much lower abundance of Hp mRNA in adipose tissue compared with liver, the contribution of adipose tissue-derived Hp to blood levels is likely well below 1%, even considering the greater tissue mass of adipose tissue (Saremi et al., 2012). Accordingly, we did not detect differences in circulating Hp concentrations in any of the BFT-related comparisons. In terms of animal health, differences might have been expected when comparing cows with different health statuses, especially for uterine infections; however, this was not the case, but also the association between Hp concentrations and metritis was reported to be weak (Hirvonen et al., 1999; Pohl et al., 2015) or even absent (Yasui et al., 2014).

3.5.1 Study limitations and other considerations

One of the limitations of this study was that DMI could not be recorded in this study, although it is likely associated with changes in BCS. However, in commercial herds such as the current study, DMI records are hardly possible. The animals in this study were from a feeding trial in which supplementation with rumen-protected Met was tested. Although the diets contained enough Met, Met was supplemented to see if more Met could affect reproduction. The Met treatment did not affect performance and health status as reported previously (Süss et al., 2019). Considering that Met acts as a methyl donor for the synthesis of carnitine, which is important for the transport of FA into mitochondria for beta-oxidation, effects of Met on lipid metabolism in dairy cows seem likely (Chandler and White, 2017). The portion of hyperketotic cows (≥ 1.2 mmol/L BHB) from the basic trial with a total of 1,709 cows was not different in the control and the Met-supplemented group (11.7% versus 12.3%); furthermore, the portion of cows with BFT < 14 mm was the same in both groups (81.9% and 82.6%, respectively; Süss et al., 2019). In contrast, in the current study with a subset of these cows, Met-treatment was significant for BHB in all

three comparisons, i.e., for BFT-ap, ΔBFT, and OC-loss/no loss. The portion of Met-supplemented and non-supplemented cows was similar in each cluster (Supplemental Figure S1). However, given these differences being limited to the ap period, when the Met supplementation was not yet started, the significance for Met in the general model that included both ap and pp values, is not considered as meaningful, and we thus refrain from further interpretation. In general, the lack of differences between the control group and the Met group on any of the variables tested could be related to the relatively late start of supplementation, i.e., only within the first 20 days after calving. (Batistel et al., 2017) recommended starting supplementation one week before calving and continuing it for at least 60 days after calving to observe differences in lactation performance.

3.6 Conclusion

The variables measured in blood confirmed that cows with greater BFT losses had increased lipolysis and ketogenesis. Cows that gained BFT did not differ from normal cows in their serum variables. Differences in milk yield were limited to the BFT-ap clusters with cows categorized as VF producing less milk than the other clusters. Non-uniform adaptive responses to lactation were particularly observed for the over-conditioned cows: OC-no loss cows produced less milk and also had increased lipolysis and ketogenesis as compared to OC-no-loss cows. This study outlines the variability in the intensity of mobilizing body fat reserves in response to the metabolic changes related to the onset of lactation in cows of different body fat content before calving. The findings derived from a relatively large number of cows from the same herd on a commercial dairy enterprise are of interest for both dairy farmers and researchers; the clustering approaches described herein provide a quick and efficient method to identify extreme groups for specific management and also for outlining further comparisons, e.g., when planning more

expensive and sophisticated analyses such as proteomics or metabolomics on selected samples.

Chapter 4 - Metabolomics

4.1 Introduction

In Chapter 3, we have confirmed that over-conditioned animals *ap* and/or animals losing a lot of condition around calving, had increased lipolysis and ketogenesis compared with cows with a normal BC and/or losing less condition. In general, we demonstrated that both clustering methods (BFT-ap and Δ-BFT) were efficient in creating relevant groups. As per the objectives of our thesis, we thus wanted to further compare these clusters to fully understand whether these groups would also be different when considering their metabolite patterns by applying a metabolomics analysis.

As previously described in Chapter 1, metabolomics is the comprehensive study of all the metabolites in a biological sample (Singh et al., 2019). Recently, the use of untargeted metabolomics revealed that most of the metabolites which increase their concentration from late pregnancy to early lactation were connected to lipid and energy metabolism, whilst the metabolites which displayed a pattern of decreasing their concentration were associated with AA metabolism (Luo et al., 2019). Identifying the metabolites and the metabolic pathways that are associated with over-conditioning during the transition period is important to fully understand the mechanisms and the reasons of over-conditioning. We thus used an untargeted metabolomics approach to provide a more comprehensive picture of the metabolism of dairy cows with different BC around parturition. The untargeted nature of the method would have permitted us to investigate whether novel compounds (not yet contained in databases, or not yet associated with the lipid or energy metabolism of dairy cows) played a significant role during the transition period (Gloaguen et al., 2017).

Moreover, since our study extends the work of Suess et al. (2019) in which the effect of dietary supplementation of Met on reproductive performance was tested in a large commercial dairy farm, we wanted to test in a subgroup of animals whether the Met

supplementation would have had some effects on the metabolism. As reported in Chapter 3, we did not observe any difference between animals receiving Met as a supplementation in terms of milk production or classical blood variables. We hypothesised that the lack of differences when considering these two elements could have been due to the relatively late start of the supplementation, i.e. starting only within the first 20 days after calving, whereas Batistel et al. (2017) recommended starting the supplementation one week before calving. However, we hypothesized that even though no effects were seen in our previous analyses, that the metabolome will differ between the control and the Met supplemented group. This hypothesis is formulated form the premise that Methionine is involved in several crucial functions in metabolism; in dairy cows Met is particularly important during the transition period, e.g., for lipoprotein synthesis in the liver, synthesis of antioxidant proteins, synthesis of immune-related proteins, and the synthesis of carnitine (Zhou et al., 2016c; Batistel et al., 2017; Chandler and White, 2017).

To test this hypothesis we set the following objectives:

(1) evaluating how the metabolome differs between animals with different BC *ap* and/or with a different BC loss around calving;

(2) studying the differences between animals receiving a control diet or a diet supplemented with Met.

4.2 Material and Methods

4.2.1 Animals

Starting from the dataset used to analyse the classical variables (already described in Chapter 3), a meaningful subset of samples to study the effects of the Met supplementation and the effects of the BC on the metabolome of dairy cows after calving had to be selected for the cost-efficient use of the complex analyses.

From the 713 cows that were previously mentioned, we further limited the subset to those cows that were inseminated before 70 DIM (i.e., not requiring any hormonal treatment), and animals considered free from endometritis (healthy) or diagnosed with a severe endometritis, as explained in the Chapter 3 and in the paper of Süss et al. (2019), resulting in the selection 184 animals. From these animals, untargeted metabolomics (LC-MS) was carried out in two batches of the serum samples that were collected at 31 ± 3 DIM and stored at -80 °C until analysis.

For our analyses, a subgroup of 177 animals was analysed to study the BC loss effect using PiMP (explained below), following the clustering showed in Chapter 3, animals were grouped according to their extent of BS and BFR loss: Extreme loss (n = 7), Moderate loss (n = 26), Small loss (n = 86), No loss (n = 44), and Gain (n = 14). Moreover, animals from Batch 1 (n = 92) were used for the preliminary evaluation of the dietary treatment, including control cows (n = 54) and the cows receiving the Met supplement group (n = 38). In contrast, all the 184 animals were used to study the interactions between BC loss, diet treatment, and uterine health status (healthy or diagnosed with endometritis) when considering the batch corrected data as explained below.

4.2.2 Metabolomics analyses

This part of the study has been done in collaboration with Glasgow Polyomics, which provided support in terms of the study design, sample preparation, and data generation and analyses.

Samples preparation

All solvents used were of the highest purity and suitable of LC-MS analysis. The samples were thawed at 4 °C and metabolites were extracted using a chloroform, methanol,

and water (1:3:1) mixture. A mix of chloroform, methanol, and water (1:3:1; 975 µL) was added to 25 µL of serum sample and extracted on a cooled (4 °C) vortex mixer for 5 min. The mixture was centrifuged for 3 min at 13,000 *g* at 4 °C, and then the supernatant was separated and stored at -80 °C until used for the LC-MS analysis.

Liquid chromatography-Mass spectrometry (LC-MS) analysis

All samples were analysed on a Thermo Scientific Q-Exactive Orbitrap mass spectrometer running in positive/negative switching mode. This was connected to a Dionex UltiMate 3000 RSLC system (Thermo Fisher Scientific, Hemel Hempstead, UK) using a ZIC-pHILIC column (150 mm x 4.6 mm, 5 µm column, Merck Sequant, Gillingham, UK). The column was maintained at 40 °C and samples were eluted with a linear gradient (20 mM $(NH4)_2CO_3$ in water, A, and acetonitrile (Rathburn Chemicals Limited, UK), B) over 26 min at a flow rate of 0.3 mL/min represented in Table 5.

Table 5*. The solvent gradient of 20 mM (NH4)2CO3 in water, A, and acetonitrile, B for the metabolomics analysis*

Time (min)	% A	% B
0	20	80
15	80	20
15	95	5
17	95	5
17	20	80
24	20	80

The injection volume was 10 µL and samples were maintained at 5 °C before injection. For the MS analysis, a Thermo Orbitrap Q-Exactive (Thermo Fisher Scientific, Hemel

Hempstead, UK) was operated in polarity switching mode and the MS settings were as follows:

- Resolution: 70,000
- AGC: 1e6
- m/z range: 70-1050
- Sheath gas: 40
- Auxiliary gas: 5
- Sweep gas: 1
- Probe temperature: 150 °C
- Capillary temperature: 320 °C

For positive mode ionisation: source voltage +3.8 kV, S-Lens RF Level 30.00, S-Lens Voltage -25.00 (V), Skimmer Voltage -15.00 (V), Inject Flatopole Offset -8.00 (V), Bent Flatapole DC -6.00 (V). For negative mode ionisation: source voltage-3.8 kV. The calibration mass range was extended to cover small metabolites by the inclusion of low-mass calibrants with the standard Thermo calmix masses (below m/z 138), butylamine ($C_4H_{11}N_1$) for positive ion electrospray ionisation (PIESI) mode (m/z 74.096426), and COF_3 for negative ion electrospray ionisation (NIESI) mode (m/z 84.9906726). To enhance calibration stability, lock-mass correction was also applied to each analytical run as shown on the next page in Table 6.

Table 6*. Lock-mass correction applied to each analytical run of the metabolomics study for both positive and negative mode*

	Positive Mode Lock masses	**Negative Mode Lock masses**
Number of lock masses	3	1
Lock mass #1 (m/z)	83.0604	89.0244
Lock mass #2 (m/z)	149.0233	/
Lock mass #3 (m/z)	445.1200	/

Samples were divided into two batches of 92 samples each, and they were run with the same machine settings.

Twenty pool samples have been created by adding 5 µL of each samples and have been analysed. These pool samples were then compared and used to assess the quality control. Moreover, six matrix blank were used to ensure that the data matrix used for statistical analysis and biological interpretation reflected the biological system being studied.

4.2.3 Data analysis

The data obtained from the LC-MS analysis were converted with the MS convert tool of Proteo Wizard (msConvertGUI, http://proteowizard.sourceforge.net/) into mzXML files for PiMP (Polyomics integrated Metabolomics Pipeline) and also into mzML files for the batch correction part. Generally, an mzXML document contains all the information of one level of MS (Deutsch, 2010). An mzML contains the same information but also carries the fragmentation data.

We analysed our data using two different approaches: one (PiMP) to analyse pairwise and combinatorial comparisons, and a batch correction pipeline to analyse more complex data.

The Polyomics integrated Metabolomics Pipeline (PiMP)

The Polyomics integrated Metabolomics Pipeline (PiMP) is a comprehensive and integrated web-enabled pipeline that offers automated and user-friendly analysis from mass spectrometry data acquisition to biological interpretations (Gloaguen et al., 2017).

The processing of metabolomics data in PiMP is made of five principal and sequential tasks: (1) project administration, (2) data upload, (3) quality control, (4) analysis parameters, and (5) data interpretation. The pipeline supports pairwise and combinatorial comparisons only.

The analysis components in PiMP are based on an R pipeline built around the XCMS software (http://metlin.scripps.edu/download/; Smith et al., 2006) for the feature detection, and mzMatch.R (Scheltema et al., 2011) for the general metabolomics pre-processing tasks, such as alignment, batch correction, and identification (Gloaguen et al., 2017).

Once the analysis is completed, the results are returned on the same web application via a PiMP-specific XML format, which can allow further analysis of the data with a new computational pipeline, if the same output schema is maintained (Gloaguen et al., 2017). The same results can also be exported into text files for further processing outside PiMP.

In our study, we used PiMP to evaluate the preliminary results of the effect of the BC loss around parturition, and the effects of the Met supplementation. It is important to underline that this approach permits analysis of only one batch at a time.

Batch correction pipeline developed by Glasgow Polyomics

In our second approach, we wanted to have a complete overview of our data, thus understanding whether there were any differences when considering the BC loss, the diet, and the uterine health status of the dairy cows. Moreover, we wanted to analyse the data of both batches at the same time. As multiple comparisons were needed, PiMP could not be

used; for this reason, a different pipeline developed by Glasgow Polyomics was used. The first step of this analysis is the batch correction, as our sample size was particularly large and had to be run in two batches. This correction is necessary because even though data are generated from the same machine, they often show different characteristics (Liu et al., 2020). Multiple approaches can be used to solve this problem, such as processing the data as a single batch and consider the variation of between-batch data as random noise, or pre-processing each batch individually, followed by an alignment of features between the features from each separate batch (Brunius et al., 2016; Liu et al., 2020).

In our study, a pipeline in R (R Core Team, https://www.R-project.org/, 2020) developed by Glasgow Polyomics was used. The used R code is provided in the Annex, and a brief description of the pipeline is provided in the next paragraph.

Similar to PiMP, the pipeline for the batch correction works combining both XCMS and mzMatch.R, and it requires a series of sequential steps, briefly described below:

- Relative Standard Deviation (RSD) filter: it calculates the relative standard deviation (the ratio of the mean to the standard deviation) for all features. Any features with an RSD lower than a predefined threshold are excluded;
- Minimum detection: the number of measured values is counted for each feature, and any of the features with less than a predefined minimum number of values is removed;
- Noise filter: it permits to reduce the level of noise and redundancy of the signals;
- Intensity filter: it is an equalization step that removes the signal intensity drift between the two different batches;
- Gap filling: it allows us to recover features that were discarded during the data pre-processing or were simply not detected;

- Batch correction: it is the final step that puts together all the filters applied before and it normalises them.

Once all the filters have been applied, the peaks are merged in a single mzXML file, ready to be processed for statistical analysis.

4.2.4 Statistical analyses

Following the PiMP pipeline, data interpretation was performed within the web platform itself. First, the peaks were annotated/identified based on mass and mass/retention time match to known standards respectively. Second, group-wise comparisons were performed to identify differences (fold-change). The fold-change is then associated with a statistical difference. Within this analysis, peaks with an adjusted p-value, calculated by applying the Benjamini-Hochberg procedure which decreases the false discovery rate, lower than 0.05 were considered significant.

When considering the other pipeline, batch corrected data were subsequently analysed using the online platform Metaboanalyst (https://www.metaboanalyst.ca), an easy-to-use online web tool that can be used to analyse a multitude of data, with metabolomics being one of them. The platform permits the performing of data analysis, data interpretation, and, eventually, integration with other omics data (Chong et al., 2019).

The first process when uploading data on Metaboanalyst is data normalization, and it is formed by three steps: row-wise standardisation, data transformation, and data scaling. In our project, data were not row-wise standardized as this step was already performed within our metabolomics data analysis pipeline (see above). However, we did transform our data via a Log10 transformation, and we applied a range scaling, by which the means were centred and divided by the value range of each feature. An example of normalization

is given in Figure 11, from which we can see the transformation of skewed data (left) into approximately conform data.

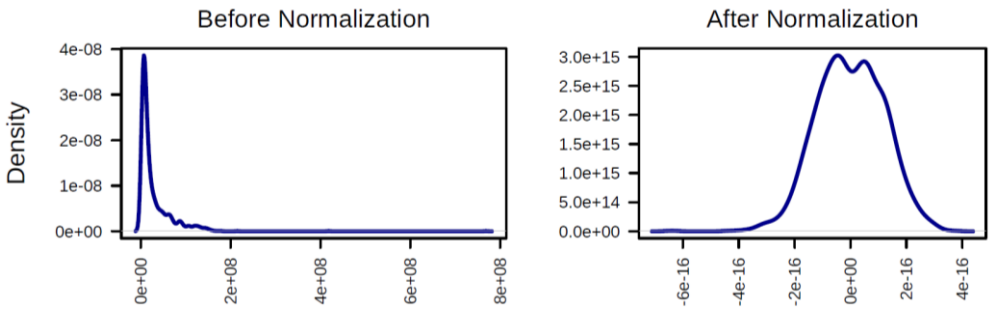


Figure 11. *Kernel density plots before(left) and after (right) normalization. The distribution of data over a continuous interval is shown. In the graphs represented herein, all the 184 samples have been normalised.*

Following the normalisation, both univariate (one-way ANOVA) and multivariate analyses (Principal Component Analysis (PCA) and Partial Least Squares-Discriminant Analysis (PLS-DA)) methods were applied. When considering the ANOVA, pairwise comparisons were also done using Fisher's least significant difference method (Fisher's LSD). Results were considered significant with a P-value < 0.05.

4.3 Result

Before heading into the results, we observed the chromatograms of both positive and negative ion mode (Figure 12).

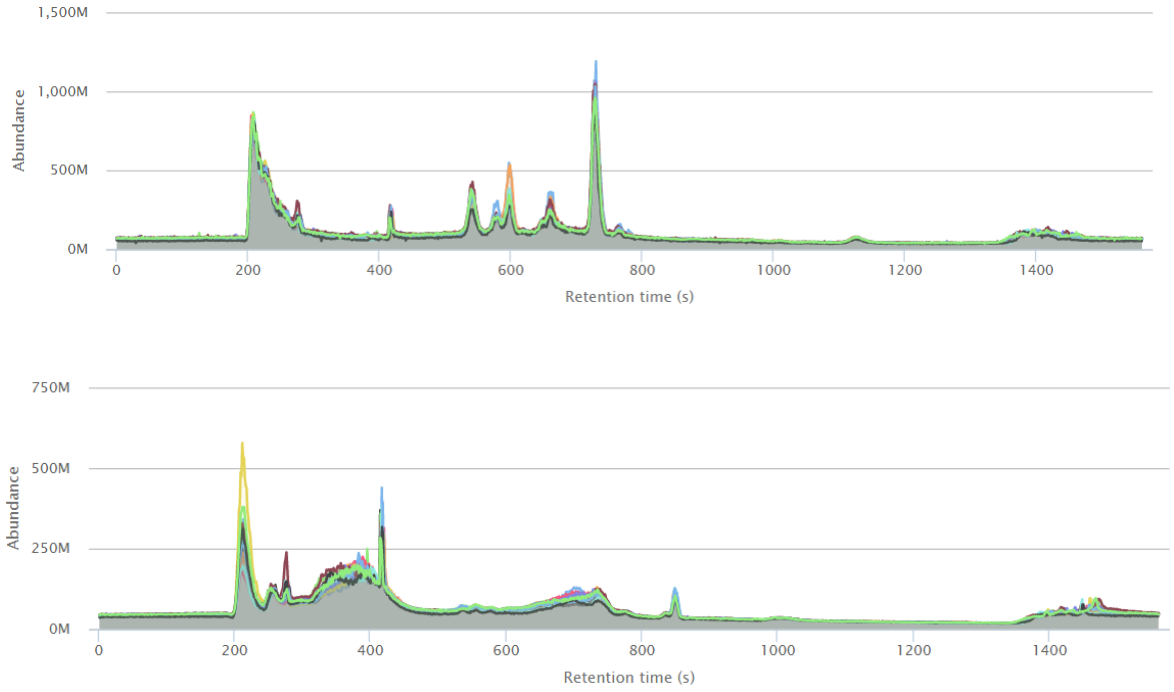


Figure 12. *Representation of two exemplar figures relative to positive ion (top) and negative ion (bottom) mode chromatograms of bovine serum. Specifically, the graphs represent one cluster (Small Loss) of Batch 1. X-axes represent the retention time (t) and the y-axes the abundance.*

4.3.1 Application of the PiMP pipeline

BC Loss

The two batches were singularly analysed on PiMP to have a preliminary overview of the quality of the analyses.

Metabolomics analysis identified 76 unique compounds matched to known standards, from the analysis of Batch 1, and 82 unique compounds for Batch 2 (full list on the Supplemental Table S4). Pairwise comparisons have been performed between the five clusters, in both batches, but without finding significant differences.

Subsequently, data from Batch 1 were exported, and evaluated using MetaboAnalyst. First, a PCA was plotted (Figure 13) as we wanted to observed trends, visualise the distribution of the cluster and possible outliers.

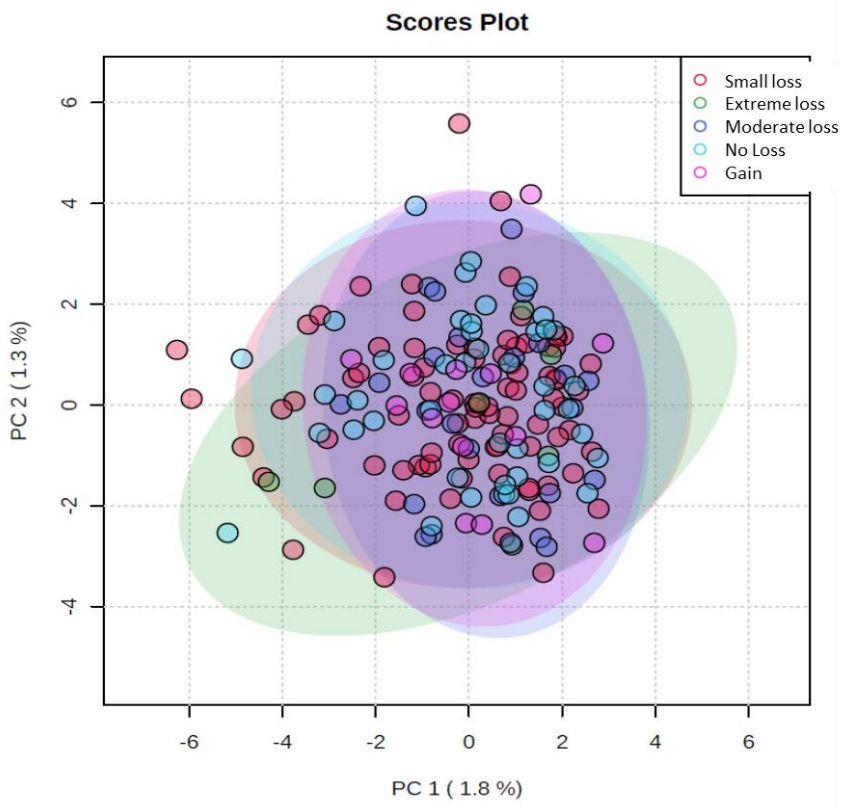
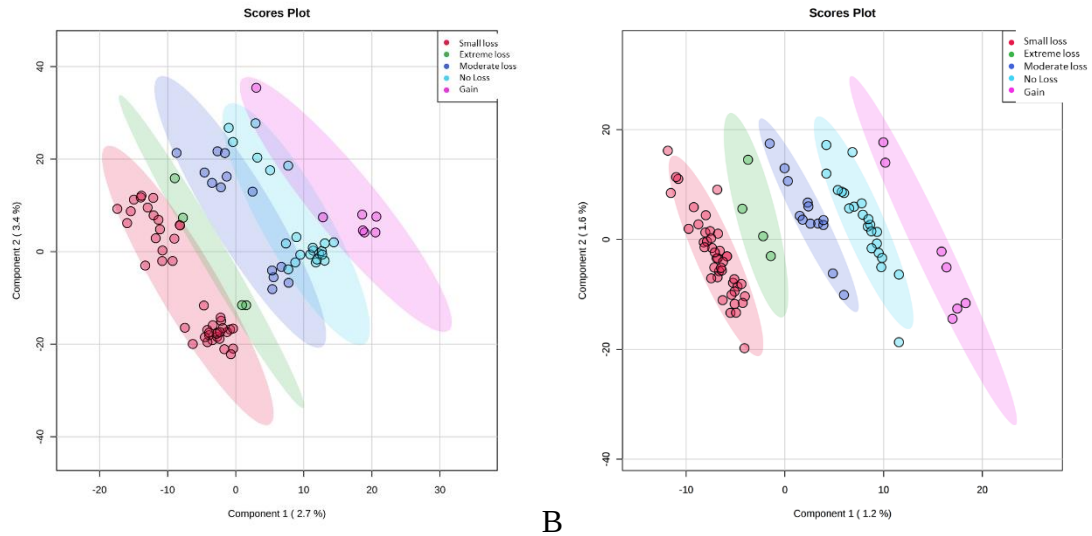


Figure 13. *Principal Component Analyses of the 92 samples of Batch 1 to study the differences between the different clusters. Each dot in the graph represents a cow in each cluster.*

The PCA did not show any separation between the clusters and also the components (PC 1 and PC 2 forming the x and y-axis, respectively, percent values) were very low, considering that 70% amongst the first two or three components are commonly considered as acceptable (Jollife and Cadima, 2016).

To further exploit our data, a PLS-DA was performed. When compared to PCA, PLS-DA has the advantage of achieving a dimensionality reduction but maintaining the awareness of class labels (Ruiz-Perez et al., 2020). As shown in Figure 14A, the five clusters were separated, however, there was a clear division in each cluster, which likely indicated that peaks could have been identified as background noise.



A B

Figure 14. *Partial Least Squares - Discriminant Analysis (PLS-DA) of the 92 samples of Batch 1. A: PLS-DA before the screening of the dataset for background noises and peaks not associable to metabolites. B: PLS-DA after the filtering of the dataset. Each dot represents a cow in each cluster.*

For further filtering the dataset, we screened the data by removing all the peaks which could have been identified as background noise as well as peaks which were not associable to metabolites; we then reapplied a PLS-DA on these screened data. In Figure 14B, the clusters were well separated, and the distribution of the samples within each cluster was more homogenous. However, the components (PC 1 and PC 2 forming the x and y-axis, respectively, percent values) were still very low. To estimate the error rate of this analysis cross-validation was performed; the respective results are presented in Figure 15.

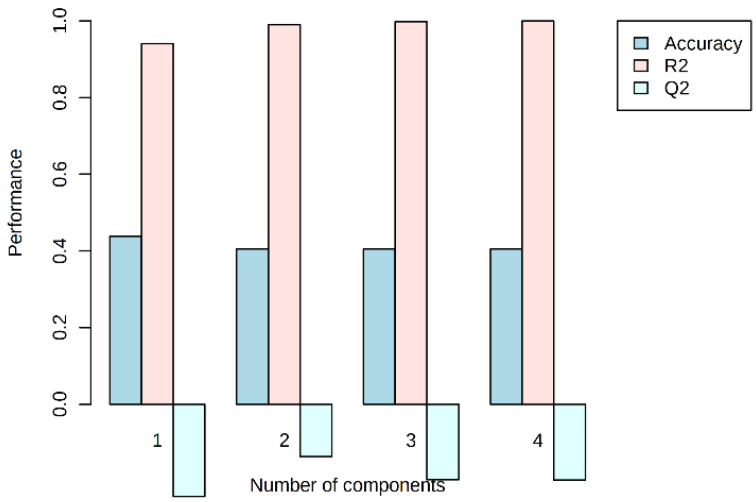


Figure 15. *Plots obtained by the cross-validation method applied on partial least squares-discriminant analysis (PLS-DA) data. The PLS-DA cross-validation data displayed a negative Q2, indicating that the model is not predictive or is overfitted.*

R2, defined as "the part of variance of the responses that is explained for the calibration model samples," is normally used to measure the quality of the prediction model. However, when using a cross-validation model, this parameter is called Q2, and represents the quality of the model prediction. Moreover, the higher the difference between R2 and Q2, the less predictive the model is (Bevilacqua and Bro, 2020).

The Q2 values which estimates the predictive ability of the model, should be > 0.5 (Golbraikh and Tropsha, 2002). A negative value of Q2, as shown in Figure 15, means that the model is not predictive or is overfitted, likely because of the use of too many variables and thus, the indicated differences between the clusters could not be considered any further.

Dietary supplementation (Con versus Met)

Batch 1 was analysed on PiMP to have a preliminary overview of the quality of the analyses and to determine whether there was a diet treatment effect in our study.

Metabolomics analysis identified 65 unique compounds matched to known standards (full list on the Supplemental Table S4). One group-wise comparison has been performed between the control and the Met supplement group, but no differences were observed.

Subsequently, data from Batch 1 were exported and evaluated using MetaboAnalyst.

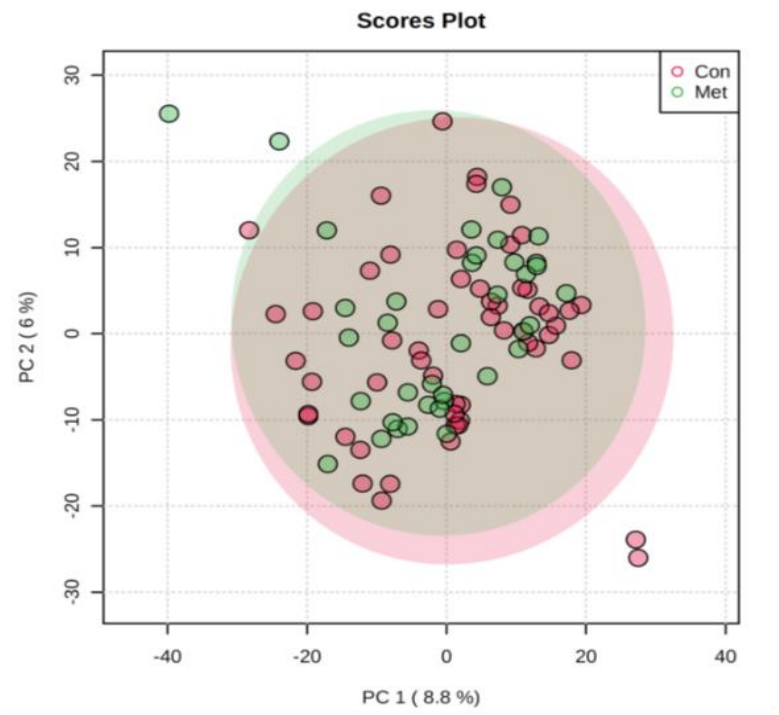


Figure 16. *Principal Component Analyses of the 92 samples of Batch 1 to study the differences between the different dietary treatments. Each dot in the graph represents a cow in each group.*

The PCA (Figure 16) did not show any separation between the groups and also the components (PC 1 and PC 2 forming the x and y-axis, respectively, percent values) were very low. As per the other analyses that were carried out, also in this case there was no clear separation of the data in the two groups, and the cross-validation results following the PLS-DA were low, i.e. negative Q2, meaning again that the model was not predictive.

4.3.2 Application of Batch corrected data pipeline

Following the preliminary results, data were batch-corrected as previously described to understand whether there were any differences when considering all the 184 samples and considering all the different factors (diet treatment, BC loss, and uterine health status). To evaluate the batch corrected, the BC loss was used as the main factor and the other two (i.e. diet treatment and health status) were considered as cofounding factors. The batch correction led us to obtain a high number of peaks (around 20,000), and due to the untargeted metabolomics methods that were used, a comprehensive assignment of them to defined metabolites was not possible. Some peaks were identified against authentic standards, defined as authentic compounds that are collected in a spectral library and subsequently used to compare with the peaks acquired from biological samples (Xiao et al., 2012). Only 73 metabolites were identified from the negative mode, and 46 metabolites from the positive mode.

An ANOVA analysis was performed, but no significant differences between groups were found. Pairwise comparisons were performed between the five clusters as well, but also yielded no differences.

Similarly, the PCA (Figure 17) did not show any separation between the clusters, and also the components (PCA1 and PCA2 forming the x and y-axis, respectively, percent values) were very low.

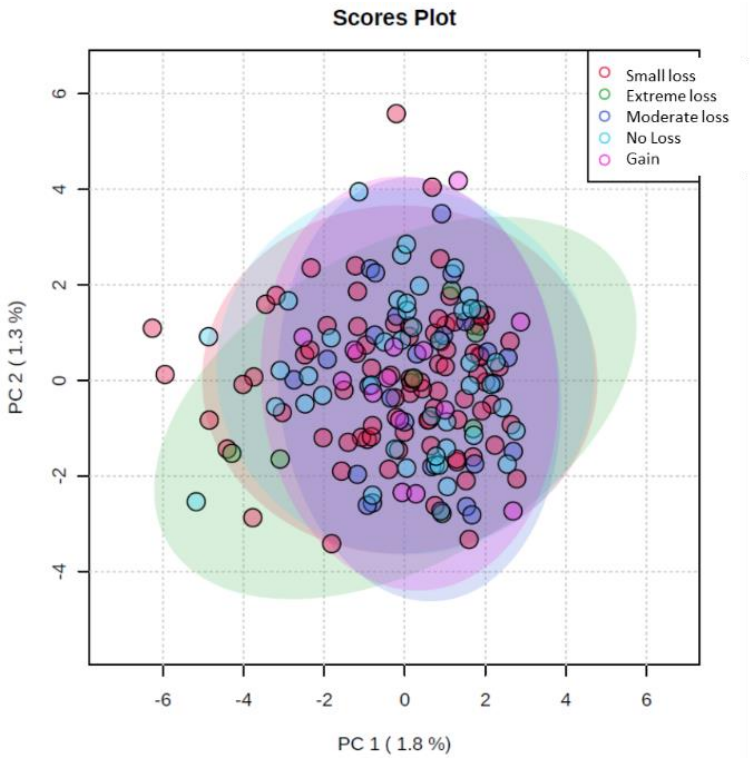


Figure 17. *Principal Component Analysis of the 184 samples to evaluate potential differences between the different clusters, with diet treatment and uterine health status as co-factors. Each dot in the graph represents a cow in each cluster.*

The PLS-DA (Figure 18) showed again no clear separation between the five clusters but the distribution of the samples within each cluster was homogenous. The cross-validation (Figure 19) showed a negative Q2 as previous results.

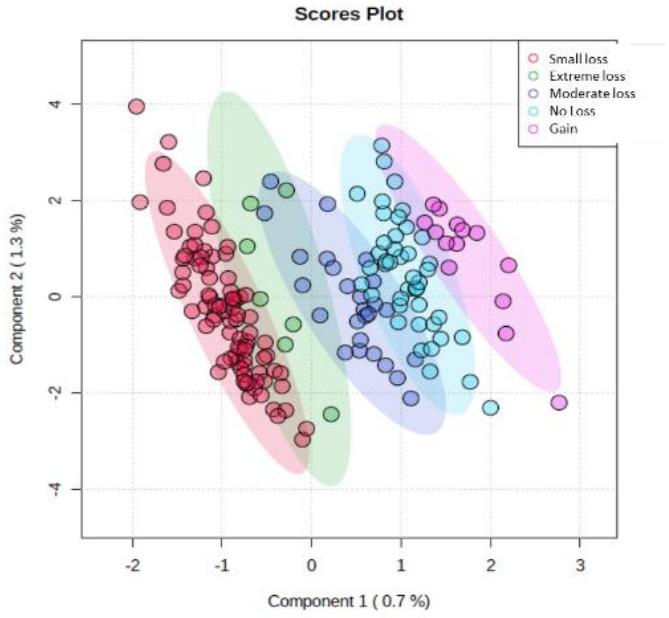


Figure 18*. Partial Least Squares - Discriminant Analysis of the 184 samples of the batch-corrected data. Each dot represents a cow in each cluster.*

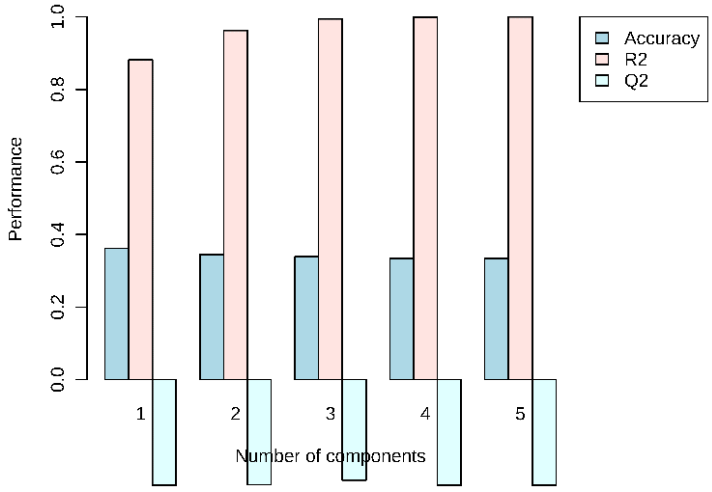


Figure 19. *Plots obtained by cross-validation method applied on partial least squares-discriminant analysis (PLS-DA) data. The PLS-DA cross-validation data displayed a negative Q2, indicating that the model is not predictive or is overfitted.*

4.4 Discussion

The combination of XCMS and mzMatch.R. XCMS (Smith et al., 2006) applied herein is the most cited pre-processing tool in the metabolomics literature (Coble and Fraga, 2014). It incorporates matched filtration, peak detection, retention time alignment, and peak matching. As for its pipeline, it first filters the data by peak detection, it then looks for common peaks between the different files and it finally uses them as standards to calculate a non-linear retention time to correct the profiles of each data file (Coble and Fraga, 2014). When the data are finally aligned, it reports the peaks into a table for statistical analysis. On the other side, mzMatch.R (Scheltema et al., 2011) is a modular, open-source, and platform-independent data processing pipeline for metabolomics data. This platform was based on the PeakML file format, which is the more common framework used between all the tools and can be thus integrated with XCMS. As for its pipeline, mzMatch.R features several tools, which can be identified as: peak extraction, matching, filtering, normalisation, derivative detection, and identification (Scheltema et al., 2011). Other pre-processing tools exist, such as MZmine (Katajamaa and Orešič, 2005), the second most cited pre-processing tool in metabolomics-related literature (Coble and

Fraga, 2014). MZmine also supports several stages of pre-processing data, including spectral filtering, peak detection, alignment, and normalization (Katajamaa and Orešič, 2005). Contrary to the previous tools, apart from open sources formats (such as mzML, mzXML), it supports proprietary formats (such as Thermo RAW).

In terms of tool performance, as reviewed by Coble and Fraga (2014), both XCMS and MZmine can provide satisfactory quantitative results for LC/MS data, and both tools require a significant level of manual input when deciding which parameters to use, how to process peak tables, and during the subsequent validation. For the present thesis, we decided to pursue the metabolomics data analyses using a combination of XCMS and mzMatch.R as it is the standard work pipeline used by the Glasgow Polyomics facility.

In Chapter 3, we have described how dairy cows with different patterns of BC loss around calving were different in terms of classical blood variables indicating that animals with greater BC loss had increased lipolysis and ketogenesis. In contrast, cows that gained BFT did not differ from normal cows in their serum variables. We thus pursued a metabolomics analysis aiming at describing their differences at the metabolome level, hypothesizing that we could differentiate the metabolite patterns between an animal with a diverse BC loss around calving. This approach has been already used in humans for differentiating metabolite patterns from metabolically healthy obese and metabolically unhealthy obese patients (Bagheri et al., 2018), and was recently also used for dairy cows (Ghaffari et al., 2020a). Moreover, Luo et al. (2019) also applied an untargeted metabolomics analysis that revealed that metabolites involved in lipid and energy metabolism were increasing their concentration from late pregnancy to early lactation, whereas the metabolites related to the metabolism of AA tended to decrease. However, in our study, we did not see any differences amongst the clusters, both when using the PiMP pipeline and the batch-corrected data. A possible explanation could be that the samples selected for analysis (as mentioned previously, around 30 days after calving), were taken

too late to observe metabolic differences, as animals might have already recovered from their negative energy balance. Moreover, albeit metabolomics provides and precise description of the metabolome in a specific time point (Singh et al., 2019), longitudinal changes with time can only be addressed if samples from several other timepoints could be considered. Luo et al. (2019), for example, collected samples 21 days before calving and on the day of calving, permitting them to have an overview on the variation of the metabolome at different timepoints.

We then focused on the Met supplementation, and also in this case, we did not see any differences when comparing the control to the Met supplemented group. We previously reported (Chapter 3) that no effects were found in terms of Met treatment when considering the classical variables, and similarly, we did not find any differences at the metabolome level. We assume that the lack of differences between the two groups could be due to the late start of the supplementation, i.e. only around 20 days after calving. Batistel et al. (2017), recommended starting a Met supplementation before calving and continuing it for at least 8 weeks after calving to observe a difference in terms of lactation performance. Moreover, due to the large number of animals involved in the study of Süss et al. (2019), recording the feed intake of each cow was impossible, and thus, we would not have the certainty that all the cows in the Met group received the same amount of the supplemented AA.

Lastly, a limitation of the study could have been its complexity. The high number of samples made the comparison between the high number of peaks obtained by each samples (around 20,000) and a comprehensive assignment of them to defined metabolites difficult. There is no automatic way of doing so, and manual processing and validation of the data would have been an almost prohibitive task, as also reported by Coble and Fraga (2014). Moreover, running multivariate analyses on our complete dataset, made of roughly

20,000 signals, with some of them identified as noise, probably contributed to the reduction of the statistical power of the model itself.

Possible future steps of this research would likely involved the analyses of multiple time points to evaluate the time effect, but also a different grouping of the animals in order to reduce the complexity of the study. A targeted approach might also be carried out to evaluate whether specific metabolites differed between the different clusters (e.g. metabolites related to the lipid metabolism).

4.5 Conclusions

The metabolomics analysis used herein showed neither differences between the BC-loss clusters nor the groups receiving the Met supplement or not. Our initial aim to describe the differences at the metabolome level of dairy cows with a different BC loss was not fulfilled by this study. As limitation, the variety of co-factors affecting our study likely contributed to creating a complex data analysis. Further studies would likely have to consider more time points within the transition period, and a more controlled study in terms of onset of the treatment as well as dosage of Met should be carried out in less animals. Moreover, a targeted approach might be useful to quantify metabolites level and determine the differences found in the pathways of interest.

Chapter 5 - Proteomics

5.1 Introduction

As reported in Chapter 4, we have not observed differences in the metabolome of dairy cows when considering BC-based grouping or the Met supplementation. Hence, when aiming at a proteomics approach, we decided to further narrow down the subgroups to animals with different BFT-ap but otherwise homogenous characteristics to avoid too many fixed factors in evaluating the multifactorial results.

As detailed in Chapter 1, proteomics is the study that aims at the characterization of the proteins present in a cell, organ, or organism in a specific time point (Wilkins et al., 1996). Due to recent technical developments, the application of proteomics has gained increasing interest in research including animal science (Ceciliani et al., 2018b; Ghaffari et al., 2020b). The proteins circulating in the blood are of interest as they carry valuable biological information (Geyer et al., 2016). Many studies have already evaluated and described the proteome of dairy cows affected with hypocalcaemia (Fan et al., 2017), endometritis (Miller et al., 2019), or intramammary infection (Kim et al., 2011). However, even though the association between different levels of BCS change, milk production and reproductive performance, as well as blood metabolites, was reported in several studies (Carvalho et al., 2014; Barletta et al., 2017; Gobikrushanth et al., 2019), the serum or plasma proteome was not/hardly investigated for its relation to BCS and BCS change. This approach can be beneficial as the identification of proteins that are associated with different BC around calving may explain the mechanisms involved in the pathophysiology of over-conditioning in dairy cows (Ghaffari et al., 2020b). Moreover, when conducting proteomics analysis, the serum or the plasma proteome are not the only targets. The characterisation of the liver proteome (Swartz et al., 2021) identified proteins involved in fatty acid metabolism, mitochondrial dysfunction, and inflammation. Similarly, the characterisation of the adipose tissue's proteome (Takiya et al., 2019) can also be very

informative when aiming at studying the metabolism. It can be considered as an endocrine organ that also produces cytokines, which can then modulate the inflammatory response (Han et al., 2017).

A factor to consider when working with proteomics, as already mentioned in Chapter 1, is the concept of high- and low-abundance proteins. The depletion of the most abundant proteins is often performed to increase the depth of proteome identification and to subsequently increase the sensitivity of the analyses (Tu et al., 2010). However, the depletion might also lead to bias as there might be cross-reactions of the antibodies used in immuno-depletion (Bellei et al., 2011) and thus compromise the informational value. However, recent bioinformatical developments have improved the use of un-depleted proteins (Amon et al., 2019). We therefore used in a pilot experiment the depleted and fractionated serum samples to enrich the spectral library used for the actual measurements. The results of depleted and un-depleted pilot samples were then combined from a regular data acquisition method ("data-dependent"). As expected, more protein groups were identified in un-depleted than in depleted/fractionated samples. To avoid the problem of missing values, we applied a "data-independent acquisition" method for the actual 30 serum samples. The comparison of non-depleted vs. depleted samples would only have yielded technical knowledge but no biological insights. With the possibility of the data-independent acquisition, we finally used un-depleted serum samples.

My objective was to determine whether the serum proteome of dairy cows that are fat *ap* was different from lean animals. To pursue this comparison, we also decided to limit the effect of possible cofactors, and thus selected animals that were all healthy and in the control group, i.e. not receiving Met supplementation.

5.2 Material and Methods

5.2.1 Animals

Starting from the dataset used for the metabolomics analysis (already described in Chapter 4), we further selected animals as we wanted to reduce the number of factors that could be confounding factors in the analysis. In particular, we wanted to have a complete overview of the BC status of the animals.

For this reason, from the 184 cows that were previously analysed in terms of metabolomics, we further selected those cows which were considered free of endometritis (as described by Suess et al., 2019) and were in the control group, thus, not receiving the Met supplementation. Second, instead of considering all the five clusters previously identified, animals representing the extreme clusters were selected, i.e. cows with a BFT < 13 mm (lean) or > 17 mm (fat) *ap* (same time point as used in Chapter 3) and having parity 2 or 3 (2^{nd} parity: n = 17; 3^{rd} parity: n = 13). A total of 30 cows (n = 16, > 17 mm; n = 14, < 13 mm) were obtained. From these animals, a proteomics analysis was carried out as described below on serum samples that were collected at 31±3 DIM and stored at -80 °C until analysis.

5.2.2 Proteomics analyses

This part of the study was done in collaboration with the Core Facility Mass Spectrometry Unit of the Institute of Biochemistry and Molecular Biology of the University of Bonn, and the Core Unit for Bioinformatics Data Analysis (CUBA) of the University of Bonn, which supported in terms of study design, sample preparation, data generation and analysis.

Library creation

A preliminary analysis of randomly selected serum samples from our study was performed for creating our library of proteins and peptides in support of future statistical analyses. To do so, 20 random serum samples were selected from our study, and they were pooled (20 µL of each sample). One pool aliquot ("Pool A") was then prepared for the LC-MS/MS analysis (see paragraphs below), whilst "Pool B" underwent a depletion procedure using the large capacity Proteominer kit (Catalog #1633007, Bio-Rad Laboratories, Hercules, CA, USA). The kit permits is formed by a highly diverse bead-based library of combinatorial peptide ligands, which simultaneously dilutes high-abundance proteins and concentrates the low-abundance ones. In particular, high-abundance proteins saturate their high-affinity ligands whilst their excess protein is washed away. On the contrary, low-abundance proteins are concentrated on their specific affinity ligand(s).
After depletion, Pool B was also prepared for the LC-MS/MS analysis (see paragraphs below).

Peptide preparations

For library generation, depleted and non-depleted protein samples were subjected to in-solution preparation of peptides on centrifugal filter units (modified from Manza et al., 2005; Masuda et al., 2008; Wiśniewski et al., 2009; Leon et al., 2013). Protein solutions were loaded onto centrifugal filter units with a 10 kDa cut-off modified PES membrane (polyether sulfone, Pall Filtersystems, Crailsheim, Germany) and reduced with 20 mM dithiothreitol at 55 °C for 30 min. Alkylation of thiol groups was done with 40 mM acrylamide for 30 min at room temperature. After another buffer exchange, 250 ng of trypsin (Promega GmbH; Walldorf, Germany) was added in 20 mM triethylammonium bicarbonate, 0.5% sodium deoxycholate (SDC) in a total volume of 50 µL. Digestion

proceeded for 10 h at 37 °C. Peptides were collected and SDC was precipitated with Trifluoroacetic acid (TFA; 2% final). The remaining SDC was removed by phase transfer with an equal volume of ethyl acetate. By using this methodology, Filter aided sample preparation, disposable centrifugal ultrafiltration units allow detergent depletion, protein digestion, and isolation of peptides which are released by proteases from undigested material. By using consecutive protein digestion with different proteases, the generation of peptide fractions with minimal overlap is enabled, which considerably increases the number of identifications and protein sequence coverage. Peptides were then pooled and dried in a vacuum concentrator, dissolved in IPG buffer (i.e. ampholyte-containing buffer) pH = 3 - 10 (GE Healthcare, Solingen, Germany), and fractionated with an OffGel device (Agilent GmbH, Waldbronn, Germany) according to the manufacturer's instructions. Dried peptide fractions were re-dissolved and desalted using ZipTip C18 tips (Thermo Fisher Scientific, Bremen, Germany). These are ready-to-use pipette-tip columns of C18 resin that permit a fast and efficient capture, concentration, desalting and elution of peptides for MALDI mass spectrometry and other methods. Two µL of the individual serum samples were prepared with iST 96x sample preparation kit (Preomics GmbH, Martinsried, Germany) according to the manufacturer's recommendations.

LC-MS/MS analysis

Peptide separation was performed on a Dionex Ultimate 3000 RSLC nano HPLC system (Dionex GmbH, Idstein, Germany). The autosampler was operated in µL-pickup mode. Peptides were dissolved in 0.1% 30 µL FA (solvent A). One µl was injected onto a C18 analytical column (200 mm length, 75 µm inner diameter, ReproSil-Pur 120 C18-AQ, 1.9 µm). Peptides were separated during a linear gradient from 5% to 35% solvent B (90% acetonitrile, 0.1% Formic acid (FA)) at 300 nL/min. The nano-HPLC was coupled online to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, Bremen,

Germany). The data-dependent acquisition was performed for library generation with a gradient length of 120 min. Ions between 350 and 1500 m/z were scanned in the Orbitrap detector every 3 seconds with a resolution of $1.2 \cdot 10^5$ (AGC target $2 \cdot 10^5$). Polysiloxane (445.12002 Da) was used for internal calibration (typical mass error ≤ 1.5 ppm). In a top-speed method, peptides were subjected to higher energy collision-induced dissociation (HCD: 1 Da isolation, threshold intensity 25000, stepped collision energy 25, 30, 35%) and fragments analysed in the Orbitrap with target $8 \cdot 10^4$ and maximum inject time 50 ms. Fragmented peptide ions were excluded from repeat analysis for 20 sec.

For data-independent acquisition (DIA) peptide separation was performed with a gradient length of 110 min. Scan parameters were adapted from (Amon et al., 2019): 40 windows of 15 Da plus 0.5 Da overlap were set covering m/z 399.5 to 1000.5. Isolated ions were fragmented with stepped HCD as above and fragments detected in the Orbitrap detector (profile mode) with a resolution of 30,000 in the range of 200-1800 m/z. AGC target was $5 \cdot 10^5$, maximum injection time 50 ms. Every 3 sec an MS1 scan was recorded (350-1500 m/z, resolution $1.2 \cdot 10^5$, target $2 \cdot 10^5$). Samples were run in three batches.

5.2.3 Data analysis

Raw data processing was performed with Spectronaut 14.10 (Biognosys AG, Schlieren, Switzerland) with a hybrid library approach that included DDA and DIA data mostly with default settings. Protein sequences were taken from Uniprot *Bos taurus* reference proteome (UP000009136, 37.882 entries) along with the MaxQuant database of common contaminants (245 entries). Fragment group-based quantifications were exported for further data analysis.

The statistical analyses of the fragment ion level data (peak-area intensity values) were performed in the R software environment (R version 3.6) (R Core Team, https://www.R-project.org/, 2020). Proteins detected by only one feature (i.e. unique

peptide) were considered as single-hit proteins and removed from the analysis. Feature level data were variance-stabilized and transformed using the VSN package (Huber et al., 2002). To estimate protein abundances following normalization, feature level data were summarized on the protein level using Tukey's median polish algorithm (Tukey, 1977).

Prior to the main statistical analysis, peptides were analysed by LC-MS/MS, then filtering and data cleaning steps were applied to clean the dataset from noise and unreliable measurements. These steps are summarised as follows:

1) Features (defined by combining the following parameters: *PeptideSequence*, *PrecursorCharge*, *FragmentIon*, *ProductCharge*) with less than 3 measurements across runs were removed,
2) Single-shot-proteins (i.e. proteins with detected via single peptide) were filtered out,
3) Feature intensities with values smaller than 0.1 were considered as missing and replaced by NA (not available),
4) Multiple measurements of the same features were tagged as individual data points.

The statistical analysis was performed on the remaining data. The dataset at this point consisted of 231 proteins, with 16739 underlying features.

Two different statistical approaches were then applied to address the missing data. In the first approach, the missing values were imputed by applying the k-nearest neighbour algorithm (KNN) (Hastie et al., 1999), implemented in the R package imputeLCMD (Package and Lazar, 2015). In the second approach, the missing values were imputed as seen before, instead, the batches were modelled using an additional fixed effect in the statistical model, as a batch effect was detected, as explained in the results section.

The differential expression analysis was carried out using the R package limma (Ritchie et al., 2015). A linear model was fitted to calculate the contrast between the two condition BFT levels, defined as > 17 mm and < 13 mm. In all analyses, the parameter

parity was added to the statistical model as fixed-effect. As explained above, in the second approach the *batch* parameter was also modelled as an additional fixed-effect.

The resulting P-values were adjusted for multiple testing and the false discovery rates (FDR) were calculated by the Benjamini-Hochberg method. The Volcano plots, heatmaps, and PCA plots were generated using ggplot2 (Wickman, 2016), ComplexHeatmap (Gu et al., 2016) and FactoMineR (Lê et al., 2008) packages, respectively.

5.3 Results

In the first approach evaluating the 30 samples, four significant proteins between the two groups (lean vs fat) were found: A5PJ69 (Serpina-10), A0A3Q1MQV3 (not known), A0A140T897 (not known), and P80195 (glycosylation-dependent cellular adhesion molecule 1). The Top 20 proteins are shown in the Annex (Supplemental Table S5).

The Volcano plots of log2 fold changes (x-axis) and their associated -log10 transformed *p*-values (y-axis) of all identified proteins in the different conditions are given in Figure 20, and the PCA is represented in Figure 21.

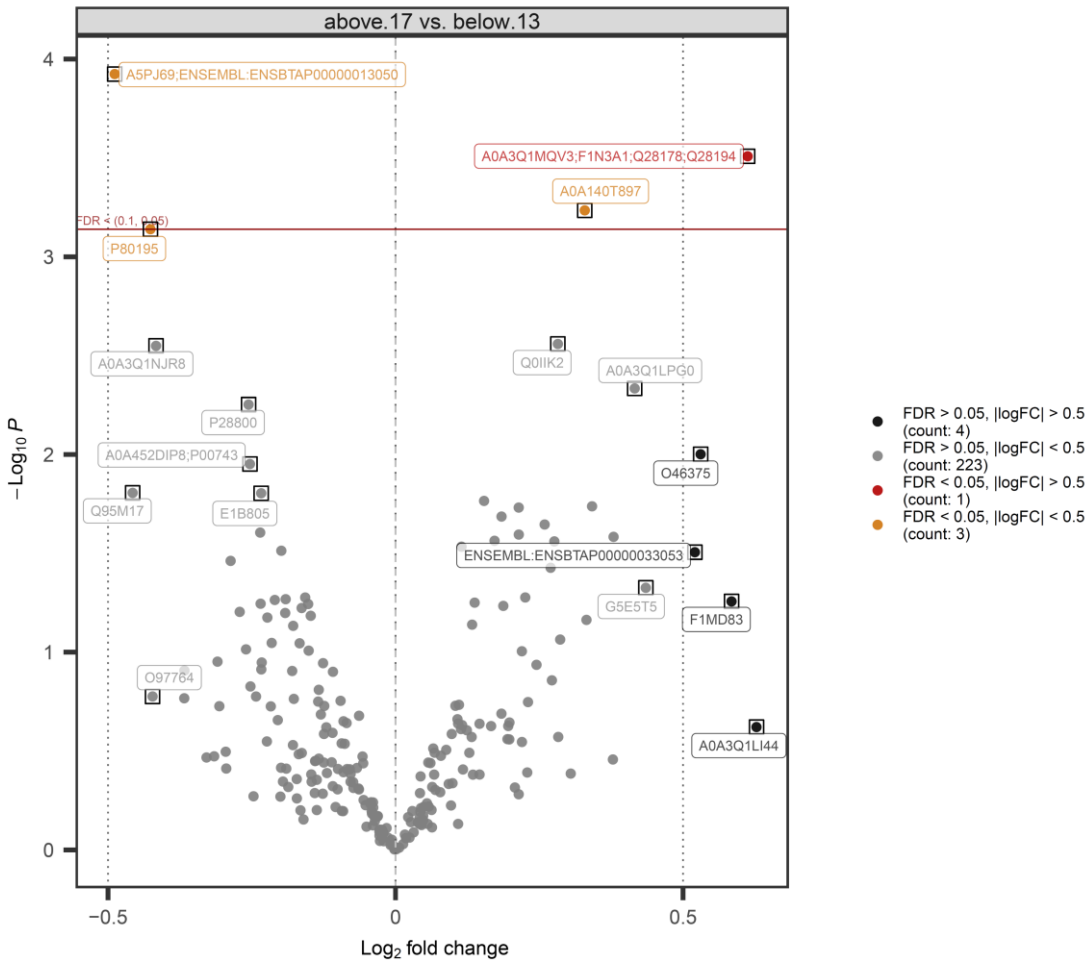


Figure 20. *Volcano plots of log2 fold changes (x-axis) and their associated -log10 transformed p-values (y-axis) of all identified proteins in the fat vs lean animals comparison. The horizontal red line represents the 5% FDR threshold on the first contrast (lean vs fat). Based on our first approach: missing values were imputed by applying the k-nearest neighbour algorithm (KNN), and parity considered in the model.*

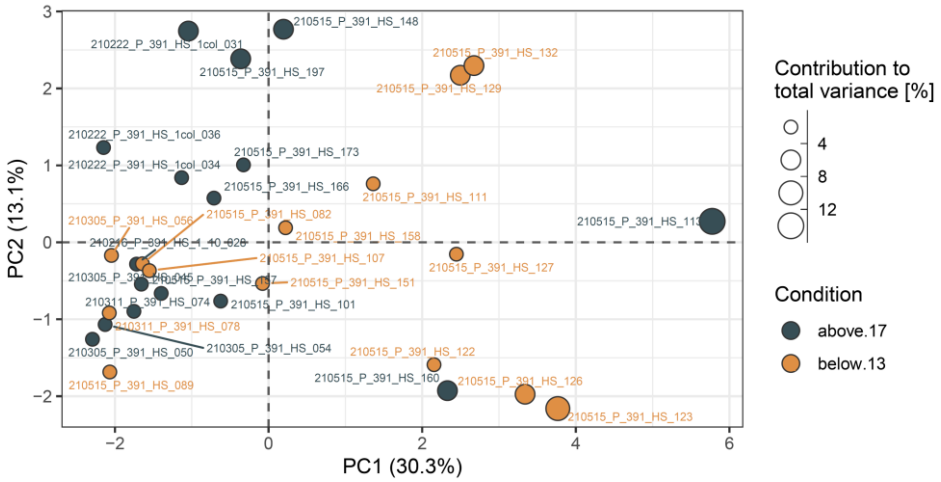


Figure 21. *Principal component analysis of the top proteins identified in our first statistical approach: missing values were imputed by applying the k-nearest neighbour algorithm (KNN), and parity considered in the model.*

However, when looking at the heatmap of the missing values (NA) in protein level data (Figure 22), the presence of a batch effect that clustered samples based on their processing date were observed.

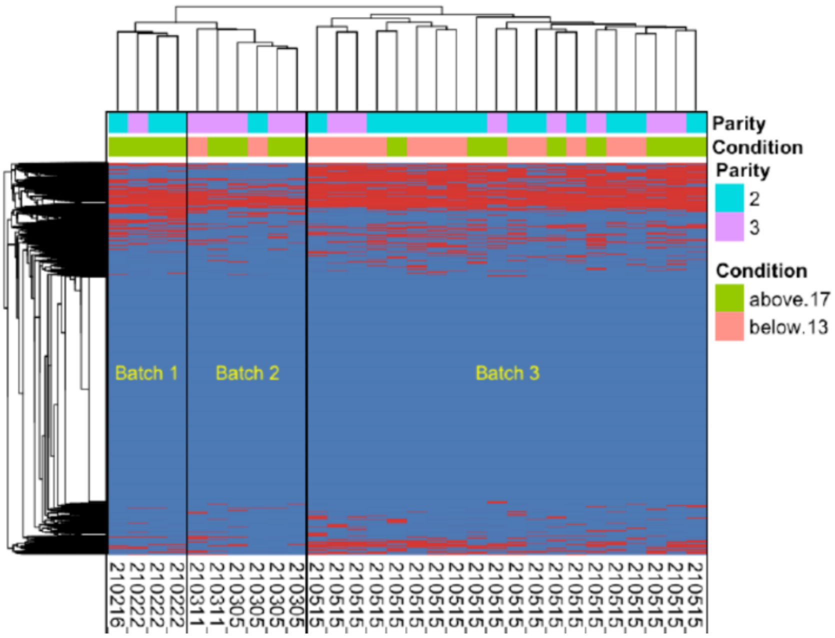


Figure 22. *Heatmap of the missing values (NA) in the 3 batches. Each Batch then defines the date of the samples analyses.*

For this reason, in the second approach, the batches were modelled as a fixed effect. With this model, no significant proteins were identified when comparing fat vs lean animals. The Volcano plots of log2 fold changes (x-axis) and their associated -log10 transformed *p*-values (y-axis) of all identified proteins in the different conditions are given in Figure 23, and the PCA is represented in Figure 24.

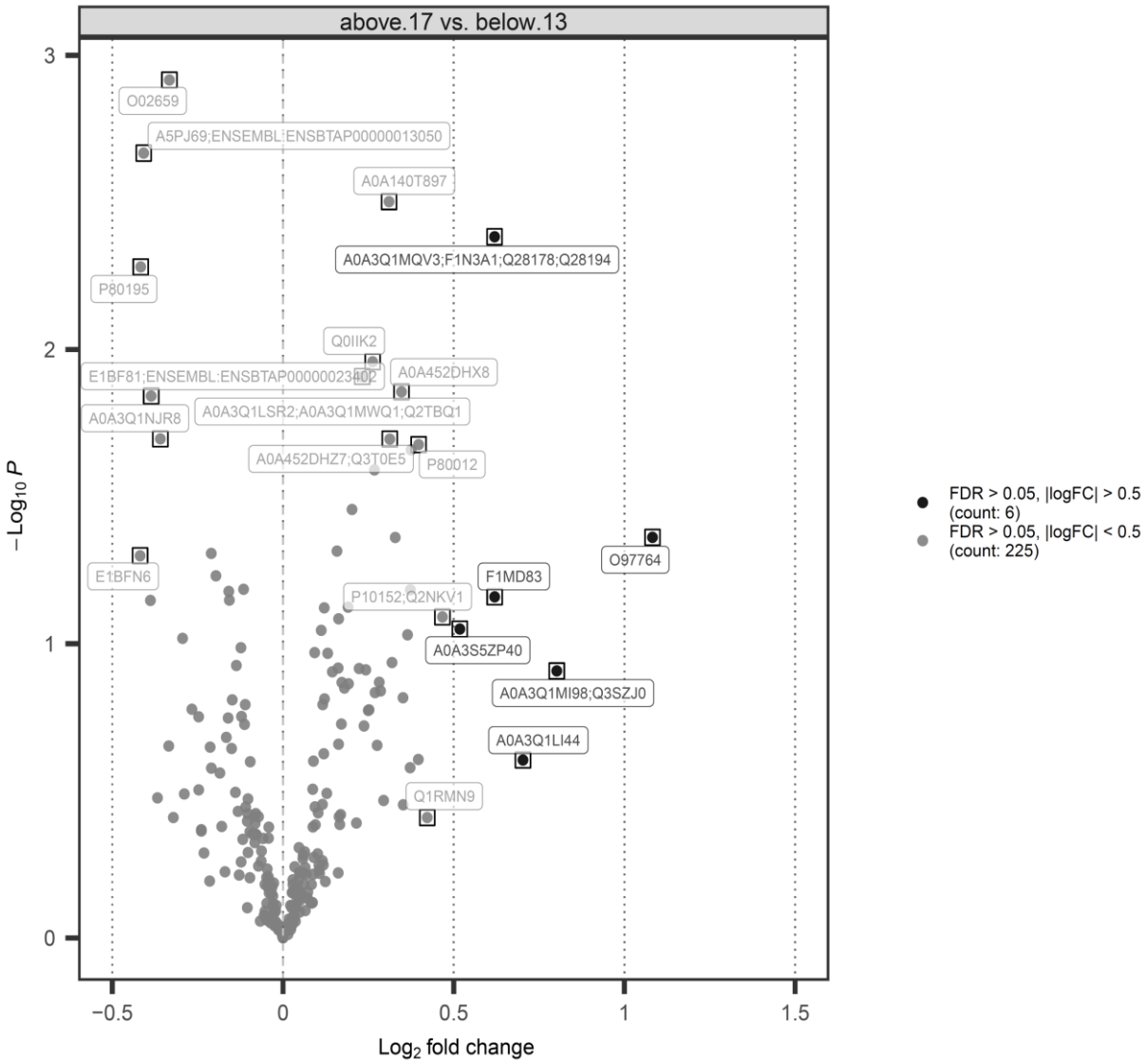


Figure 23. *Volcano plots of log2 fold changes (x-axis) and their associated -log10 transformed p-values (y-axis) of all identified proteins in the fat vs lean animals comparison . Based on our second approach: missing values were imputed by applying the k-nearest neighbour algorithm (KNN), parity and batch considered in the model.*

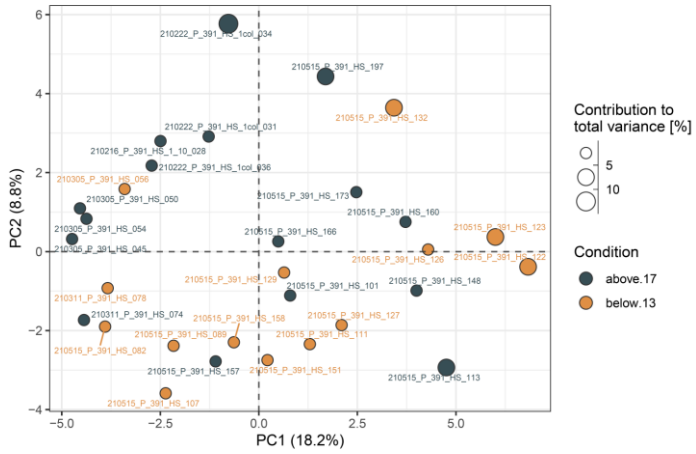


Figure 24. *Principal component analysis (PCA) of the top proteins identified in our second approach: missing values were imputed by applying the k-nearest neighbour algorithm (KNN), parity and batch considered in the model.*

However, from the top 5 proteins identified in this second approach, having a $-\log_{10} P > 2$, four were also significant in the first approach without batch correction: A5PJ69 (Serpina-10), A0A140T897 (Albumin), A0A3Q1MQV3 (Thrombospondin-1), P80195 (glycosylation-dependent cellular adhesion molecule 1 (Glycam-1)). O02659 (Mannose-binding protein C*)* was the additional protein. Specifically, the Mannose binding protein, Serpina-10 and Glycam-1 had a higher expression in the above 17 mm group compared to the below 13 mm group ($\log_2$ FC = - 0.33, - 0.41, and – 0.42, respectively). On the contrary, albumin and Thrombospondin-1 had a higher expression in the below 13 mm group ($\log_2$ FC = 0.31 and 0.62, respectively). The list of the top 20 proteins is shown in the Annex (Supplemental Table S6).

5.4 Discussion

In Chapter 3, the different patterns of BC loss around calving and the metabolic differences of dairy cows with a different BFT-ap were described by assessing some "classical" blood variables. Accordingly, greater BFT-ap was associated with increased lipolysis and ketogenesis. Evaluating the metabolome in *pp* samples of dairy cows with different BC (Chapter 4), no differences were identified. Too many factors might have compromised the informational value of the statistical models applied and it was thus decided to focus the proteomics analysis exclusively on the BC *ap*. We aimed at identifying important proteins associated with over-conditioning around calving to further understand the metabolism of dairy cows. Surprisingly, we found only 5 proteins that differed between both groups when the batch effect was not considered. When batches were considered, these proteins were no longer significant but remained as top proteins according to their FDR-values.

The mannose-binding protein C is a calcium-dependent lectin involved in the immune response; specifically, it is an acute-phase protein that has the role to activate the lectin complement pathway, and it is involved in host defence pathways (Ng et al., 1996). In dairy cows, a higher concentration of this protein could be associated to heat stress as it might activate blood coagulation (Min et al., 2016). In our study, the concentration of this protein was higher in the fat animals *ap*, however, in the literature there is no explanation why this protein may differ in expression between fat and lean dairy cows. Moreover, as the animals were managed in the same farm, a situation of heat stress would have likely affected both groups.

SERPINA10, a member of the serpin family, was reported to have a role in blood coagulation as well since it inhibits the activated factors Z and XI in blood coagulation (Law et al., 2006). However, there are no current papers reporting the effect of this protein in dairy cows or cattle, and thus, its relevance is still unknown.

The Glycosylation-dependent cell adhesion molecule 1 (Glycam-1) is an abundant protein in the plasma of dairy cows (Miller et al., 2019) and it likely inhibiting the process of cell adhesion (Choe et al., 2010). Its plasma concentration tends to be higher in animals with severe endometritis, as reported by Miller et al., (2019). In our proteomics study, we did not have animals diagnosed with endometritis, however, we have observed that Glycam-1 tended to be more abundant in the fat animals, which might indicate a higher status of inflammation in dairy cows, resulting from a shift in the production of adipokines towards a pro-inflammatory profile in over-conditioned cows (Alharthi et al., 2018; Depreester et al., 2018). Moreover, as seen in Chapter 3, the greater blood concentrations of NEFA and BHB in over-conditioned cows compared to normal or lean ones, underline an increased ketogenetic status, which can be associated with an increased level of inflammation ante and post-partum (Abuajamieh et al., 2016).

The PCA plots (Figures 21 and 24) confirmed that there was no separation of the two groups. This might be retrospectively explained by the time the samples were collected. As discussed already for the metabolome, the animals had already undergone and largely completed the process of metabolic adaptation regarding the body condition around 30 d pp, and thus potential preceding differences might have levelled off. In the study of (Ghaffari et al., 2020b) for example, samples were collected -49 d before parturition, and +7 and +21 d after parturition. In this way, a time effect could be captured and the variation of the proteome between the time points and the groups, as well as the interaction thereof was analysed. In the present proteomics analyses, measuring the samples ante and post-partum might have yielded significance and might have permitted us to evaluate how fat and lean animals cope with the transition period.

Finally, the use of a DIA approach permitted us to detect a higher number of peptides compared to a data dependent acquisition (in our study only use for the library creation), as the first method interrogates all the peptides within a selected m/z windows (e.g. 40 windows of 15 Da plus 0.5 Da overlap were set covering m/z 399.5 to 1000.5 in our study), which contains normally around the 90% of the peptides (Pino et al., 2020). Whilst a data dependent acquisition would focus on a fixed number of the most abundant peptides which would be then selected for a second stage tandem mass spectrometry (MS/MS) (Li et al., 2020). For this reason, using a data independent acquisition permitted to fragmentate all peptides which were defined within our m/z window allowing a more complete and accurate analysis of the proteins in our study, which was also important to detect the less abundant proteins, thus increasing the dynamic range of our analysis.

5.5 Conclusion

Our initial aim to describe the serum proteome of dairy cows with different BC status *ap* was thus reached, but differences between the groups could not be detected in this

study. In this proteomics approach, selecting just one time point for the analysis must retrospectively be considered as a limitation.

Chapter 6 – General discussion and Conclusions

The overall aim of this thesis was to investigate the relation between body fat and the metabolic adaptation which occurs during the transition period. For this, three studies were carried out (classical variables, metabolomics, and proteomics). The first study aimed to characterize the inter-individual variation in relation to pre-calving BFT and subsequent BFT loss during early lactation in a large dairy herd, by individuating clusters of animals with different BFT-*ap* or ΔBFT and by evaluating them in terms of milk production, health condition, and circulating concentration of metabolites (NEFA, BHB), metabolic hormones (leptin and adiponectin), and an inflammatory marker (haptoglobin) both *ap* and *pp*. The second study was conducted to characterize the metabolome of a subgroup of cows to assess the metabolic differences of animals with a different BC loss or receiving a different diet, i.e. control diet or supplemented with Met. In the third and last study, a proteomics approach was used to have a complete overview of the proteome of dairy cows selected as fat or lean *ap* but receiving the same diet and being all healthy. It was hypothesized that these studies could extend the knowledge about the biological variation in body condition around the transition period. Moreover, we expected to have a deeper understanding of the biological processes when looking at the OMICs level.

From our first approach, the present thesis characterized the relationship between *ap* BFT and the subsequent loss of BFT during early lactation in a large cohort of dairy cows. It is known from the literature that the extent of BC loss during the transition period varies between individual animals, but it is usually positively associated with the *ap* BC of dairy cows (Weber et al., 2013). In our study, we did not observe any difference when considering the milk production of animals with different BC loss around calving, but we have identified differences when considering the BFT-*ap*, as it was observed that VF tended to produce less milk in comparison to the other clusters. Gobikrushanth et al. (2019) found no association between *pp* BC and milk production but the peak yield as well as the

305-d milk yield were greater in cows with extreme losses compared to cows gaining BC. Roche et al. (2007c) instead, found a positive relationship between BCS loss (between calving and highest value of BCS) and milk yield. Moreover, in the present study, greater concentrations of NEFA and BHB were observed in over-conditioned animals *ap* and in dairy cows with greater BC losses. This is in line with the literature as it is well documented that increased lipolysis and the resulting release of fatty acids from the adipose tissue into the bloodstream, may exceed the hepatic capacity for fatty acid oxidation, thus leading to increased ketogenesis and fatty liver (McFadden, 2020; Ghaffari et al., 2021), which in turn may result in metabolic disorders (Bernabucci et al., 2005; Roche et al., 2009; Rathbun et al., 2017).

In the present study, the leptin concentrations in serum changed throughout the transition period as described earlier (Reist et al., 2003; Kokkonen et al., 2005; Schuh et al., 2019), i.e., the concentration decreased after calving in all clusters. This is connected to the onset of the NEB and a decrease in feed intake (Chilliard et al. 2005). It was also confirmed that very fat cows *ap* had greater leptin concentrations, thus underpinning the positive relationship between adiposity and leptinaemia (Chilliard et al. 2005). Similarly, the adiponectin concentration also decreased after calving in all clusters, indicating reduced insulin sensitivity; however, in contrast to leptin, there were no differences between the BC *ap* clusters. Considering the clusters made with respect to BC loss, animals with greater BC losses had lesser concentrations of adiponectin both *ap* and *pp*. Thus, BC loss seems to be closer related to the adiponectin concentration than the BC status *ap* (Singh et al., 2014; De Koster et al., 2017). Finally, when considering Hp, we haven't seen any differences among the clusters.

In the second approach, the aim was to characterise the metabolome of dairy cows clustered by their BC loss around calving, but also considering their diet (i.e. control diet vs Met supplemented). From the previous approach, different patterns of BC loss were

identified and differences were observed when considering selected blood variables that are related to lipolysis and ketogenesis. When applying a metabolomics approach, identifying the differences at the metabolome levels of the different clusters was our main goal. In recent studies (Ghaffari et al., 2020a; Luo et al., 2019), metabolomics was already used to differentiate metabolite patterns in animals with different BC. Using a targeted approach, Ghaffari et al. (2020a) were able to differentiate within a cohort of over-conditioned cows subgroups that were metabolically healthy or unhealthy. Luo et al. (2019) described the most important metabolic patterns during late pregnancy and early lactation. We also aimed at identifying such differences, however, in this thesis project no differences were identified between the clusters. A possible limitation could have been the selection of only one time point for the analysis. Besides being unable to follow longitudinal changes (which was not our goal), 30 days after calving might have been too late for identifying differences between the clusters since most of the metabolic adaptations to lactation could have already been accomplished at that time. A possible alternative, could have been to select multiple time points, as done by the works of Luo et al. 2019 and Ghaffari et al. (2020a), and observe the variation of the metabolome in the difference time point, e.g. a time point *ap*, at calving, and *pp*. However, when aiming to study physiological mechanisms underlying potential treatment effects through metabolomics on a large number of samples, selecting specific time points is desirable for a cost-efficient use of the complex analyses.

We then focused on the Met supplementation, but also in this case, no differences were observed when comparing the animal groups the Met supplemented diet or the control diet. We hypothesise that the lack of differences between the two groups could be due to the late start of the supplementation, i.e. only around 20 days after calving. Batistel et al. (2017), reported that a Met supplementation should start before calving and be continued for at least 8 weeks after calving to observe differences in milk production. Moreover, due to the

large number of animals involved in the study of Süss et al. (2019), the individual feed intake was not recorded recording, and thus, we would not have the certainty that all the cows in the Met group received the same amount of the supplemented amino acid. In support of this, no differences were observed between the two groups when comparing the (semi-quantitatively assessed) concentrations of Met from the metabolomic profiles in serum (data not shown). Finally, it is important to underline that the animals in the control group were not receiving a Met-deficient diet, and thus the additional supply of Met in the treatment group might have been unable to elicit additional effects. Both from the study of Süss et al. (2019) and the present study, no differences were observed nor in terms of fertility, nor in terms of milk production.

In the last approach, the aim was to identify important proteins associated with the over-conditioning around calving to further understand the metabolism of dairy cows. Deviating from what was done in the metabolomics approach, we further limited the number of the animals in the study to reduce the number of factors to consider in the statistical models to have appropriate statistical power. Thus, the analysis was focussed on selected dairy cows to evaluate exclusively the proteins associated with an over-conditioning around parturition. However, no proteins were significant, and the top identified proteins were the mannose-binding protein C, SERPINA10, and the Glycam-1. The former is a calcium-dependent lectin involved in the immune response which can be considered as an acute-phase protein. Its role is to activate the lectin complement pathway and it is involved host defence pathways (Ng et al., 1996). The SERPINA10 is a member of the serpin family, and it is reported to have a role in blood coagulation (Law et al., 2006), however, no relevance has been found in dairy cows up to the present time. Finally the Glycam-1 is an abundant protein in the plasma of dairy cows and it tends to have greater concentration in animals with severe endometritis (Miller et al., 2019). However, in the current proteomics study, all animals were diagnosed as free from endometritis.

Similar to the metabolomics study, also in this approach the chosen time point might have been not ideal when aiming at describing the BC loss. Selected a series of time points as done by Ghaffari et al. (2020b) might have been a better strategy, but again, the selection of one specific time point was necessary for a cost-efficient use of the complex analysis.

Moreover, with the development of new cutting-edge methodologies, future research should also focus on the phospho-proteomics of the adipose tissue in dairy cows. In fact, as also reported in the work of Daddam et al. (2021), few studies have reported the difference at the phospho-proteome in the adipose tissue. This analysis might in fact permit to have a better overview on the lipid metabolism and the regulation of protein related specifically to the adipose tissue. This new approach, together with the integration of different omics techniques, such as metabolomics, lipidomics, proteomics and possibly transcriptomics, will certainly help to produce a complete overview on the changes in the metabolomics of dairy cows and would permit to further evaluate the metabolic variation around the transition period.

Conclusions

The body condition of dairy cows is known to be related to animal health and performance. In this study, we aimed to find out more about the metabolic differences that could be present between animals with different BC *ap* or with a different BC loss. To do so, three different approaches have been used to study the metabolism at different levels: the first aiming to study "classical variables," the second aiming to study the metabolome, and the third one aiming to study the proteome. Moreover, contrary to what has been done so far, i.e. using an arbitrary threshold to classify dairy cows based on their body condition, a new approach based on machine learning was used. The clustering approach used in this thesis represents a quick and efficient method to identify clusters of similar animals and

might be used at the farm level to quickly identify extreme groups for specific management and also when planning new experiments.

This thesis confirmed that over-conditioned dairy cows and/or with a greater BC loss around calving had increased lipolysis and ketogenesis, but no differences were identified between normal loss and animals which gained condition. Milk production differences were not consistent throughout the study, meaning that there is still a great individual variability between dairy cows when adapting to the onset of a lactation.

The metabolomics and proteomics studies did not yield the results we were expecting, and thus, the initial aim of describing the difference of both the metabolome and the proteome of dairy cows with different BC was not fulfilled. The complex data analysis, the single time point analysed, and the impossibility of having a more controlled study in terms of feed intake and Met supplementation could be the main reason for the obtained results. However, as on one hand a study with a controlled number of animals might be easier to control, large animal studies permit to have a wider amount of data and might be more representative of a real farm situation. Both the approaches have their own advantages and disadvantages, and their aspects should be considered when aiming for future studies.

Chapter 7 - Summary

The transition from late pregnancy to early lactation is one of the most critical times during a dairy cows' life. The key points of this period are: drastic changes in nutrient balance, metabolic and endocrine shifts related to parturition, and the rapidly increasing milk production. To support the latter, dairy cows need to mobilize body reserves from adipose tissues; the extent of this mobilization varies between animals but is commonly more pronounced in cows that are over-conditioned at calving. This may then result in a greater risk of developing metabolic disorders, such as ketosis, and impaired fertility. Moreover, not only the level of body condition (BC) around calving is important, but also the magnitude of BC loss from pregnancy to lactation has to be considered when looking at the metabolic health of dairy cows. An excessive loss of body energy reserves has also been associated with impaired immune function, leading to an increased risk of infectious diseases.

Thus, the aims of the present thesis were (1) to characterize the variation in pre-calving back fat thickness (BFT, recorded by ultrasound, reflects body condition) and subsequent BFT loss during early lactation in a large dairy herd by relating the variation observed to milk production, health condition, and selected blood variables; (2) to perform an untargeted metabolomics analysis to characterize the metabolome in selected subgroups differing in body condition loss, health status and in dietary methionine supply; and finally (3), to perform a proteomics analysis on further selected subgroups of animals receiving the same diet and being all classified as healthy to determine the differences at the proteome level between fat and lean dairy cows ante partum (*ap*).

In the first study (Chapter 3), we aimed at characterizing the inter-individual variation in the relationship between *ap* BFT and subsequent BFT loss during early lactation in a large dairy herd, and to compare milk yield, uterine health (mainly prevalence of endometritis and hyperketonaemia), and metabolic and inflammatory status.

For the latter, the concentration of non-esterified fatty acids (NEFA), ß-hydroxybutyrate (BHB), leptin, adiponectin, and haptoglobin (Hp) were measured to provide information about ketogenesis (BHB), lipolysis (NEFA), regulation of energy metabolism (the adipokines adiponectin and leptin), and inflammation (Hp as positive acute phase-protein). For doing so, 713 animals were selected from a previous study performed in a large commercial dairy farm comprising 1,709 dairy cows in total. Blood samples were collected at the same time as the BFT measurements (25 ± 10 days *ap* and 30 ± 3 days *post partum* (*pp*) [means ± SD]). K-means (unsupervised machine learning algorithm) clustering was performed: one approach considered only the BFT-*ap*, and the second one the BFT-loss between *ap* and *pp*. Five clusters were obtained and validated for both approaches. Results from the blood variables confirmed that cows with a greater body condition loss also had elevated lipolysis and ketogenesis as compared to cows with lesser losses; animals gaining condition did not differ from normal cows in this respect. Furthermore, over-conditioned cows *ap* had lesser milk yield compared to the other cows, whilst there were no differences in milk production when comparing animals with different body condition loss. Expectedly, the leptin concentration was also greater in over-conditioned cows with greater losses. To conclude, this study described the variability in the intensity of body fat mobilization in response to the metabolic changes related to the onset of lactation in cows with a different body condition before calving.

In the second study (Chapter 4), we aimed at (1) evaluating the metabolome of dairy cows with different body condition loss around calving and being free of endometritis, and (2) studying the differences between animals receiving a standard control diet versus a diet with an extra supplementation of Methionine. To achieve this, a further subgroup was selected from the previous 713 animals, including only those cows that were inseminated before 70 days in milk and were diagnosed as free from endometritis, i.e. 184 animals entering the study. From these, an untargeted metabolomics (LC-MS) analysis was

performed in serum samples collected at 30 ± 3 days after calving. Cows were clustered as previously reported (Chapter 3), to study the body condition loss effect via the Polyomics integrated Metabolomics Pipeline (PiMP). Subsequently, the results from the analysis of Batch 1 (n = 92) and Batch 2 (n = 92) were analysed individually for a preliminary evaluation of the diet treatment, by comparing the control group (n = 54 and n = 52 for Batch 1 and 2, respectively) against the Met supplemented group (n = 38 and n = 40, for Batch 1 and 2, respectively). Furthermore, all the animals were used to analyse the interactions between BC loss, diet treatment, and uterine health status (healthy vs endometritis) via supervised and unsupervised analyses. No differences were detectable when comparing the body condition loss clusters and also for the nor the Met supplemented group versus the control group. The possible reasons for this negative finding might be related to the time of sampling but might also have been due to the complexity of the comparisons involving many co-variables.

In the third and last study (Chapter 5), we aimed at determining whether the plasma proteomes of dairy cows that are over-conditioned *ap* differed from lean animals. For this, a further subgroup of 30 dairy cows was selected, comprising animals exclusively from the control group that were all considered healthy (i.e. free from endometritis) but were classified as being either over-conditioned or lean *ap*. From these animals, a proteomics analysis was carried out on serum samples collected at 30 ± 3 days after calving. Also in the comparison, no differences between the two groups were identifiable, even though the complexity of the design was substantially reduced..

In summary, the results obtained from the analyses of the classical variables (Chapter 3) largely confirmed the literature, whilst the OMICs results (Chapter 4 and 5) could not broaden the current understanding about the relationship between body condition, fat mobilization and metabolism.

Appendix A: Supplementary materials

A.1 R code

Code used in R by the Glasgow Polyomics facility for the Pipeline of the batch correction. The script presented below is currently set up for the Negative mode. The batch correction for the positive mode is not presented as the processing is the same, the only variation is the file referencing.

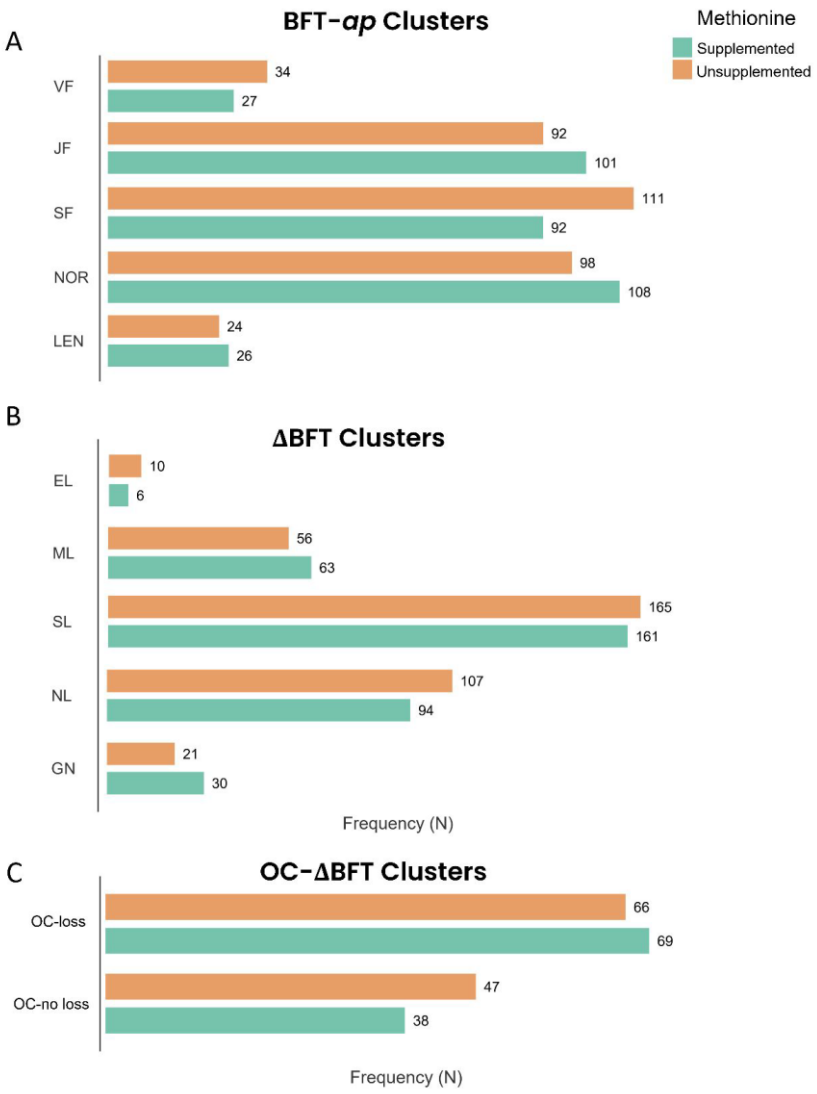
```
##Test
##HOMAGE pipeline
library(rJava)
library(mzmatch.R)
mzmatch.init(version.1=FALSE, memorysize=80000)

library(gptk)
library(outliers)
library(XML)
source("/home/gb174a/gpModel.R")
source("/home/gb174a/batchSetup.R")
setwd("/home/gb174a/RubenBatchCorrection/mzMLNegative")


##"groups" is an R list. If your files are in individual directories you can create it using:

library(tools)
groups <- list()
groups$pooled <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/pooled"))))
groups$mtxblk <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/mtxblk"))))
groups$Std1 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Std1"))))
groups$Std2 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Std2"))))
groups$Std3 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Std3"))))
groups$NoCond <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/NoCond"))))
groups$Cond0 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Cond0"))))
```

```
groups$Cond1 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Cond1"))))
groups$Cond2 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Cond2"))))
groups$Cond3 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Cond3"))))
groups$Cond4 <-
unique(file_path_sans_ext(basename(dir(path="/home/gb174a/RubenBatchCorrection/mz
MLNegative/Cond4"))))


mzMLpath <- getwd()
files <- dir(mzMLpath,full.names=TRUE,pattern="\\.mzML$",recursive=TRUE)
show(files)

inputFile = "combined_highintensity_gapfilled.peakml"
outputFile = "combined_highintensity_gapfilled_corrected.peakml"
sampleTypesToKeep <- names(groups[!names(groups) %in% "mtxblk"])

mzMLfiles <- dir(full.names=TRUE,pattern="\\.mzML$",recursive=TRUE)
outputfiles <- paste(sub(".mzML","",mzMLfiles),".peakml",sep="")


xset <- xcmsSet(mzMLfiles, method='centWave', ppm=2, peakwidth=c(5,100),
snthresh=3,
            prefilter=c(3,1000), integrate=1, mzdiff=0.001, verbose.columns=TRUE,
            fitgauss=FALSE, nSlaves=16)
##xset2 <- retcor(xset, method='obiwarp',profStep=0.01)


xsets <- split (xset,xset@filepaths)

peakMLparallel <- function(x)
{
library(mzmatch.R)
mzmatch.init (version.1 = FALSE)
xset <- xsets[[x]]
PeakML.xcms.write.SingleMeasurement
(xset=xset,filename=outputfiles[x],ionisation="detect",addscans=20,writeRejected=FALS
E,ApodisationFilter=TRUE)
}

if (length(xsets)==length(mzMLfiles))
{

cl <- makeCluster (8, type="SOCK")
clusterExport (cl,varlist=c("xsets","outputfiles"))
system.time(clusterApply(cl,1:length(outputfiles),peakMLparallel))
stopCluster(cl)
```

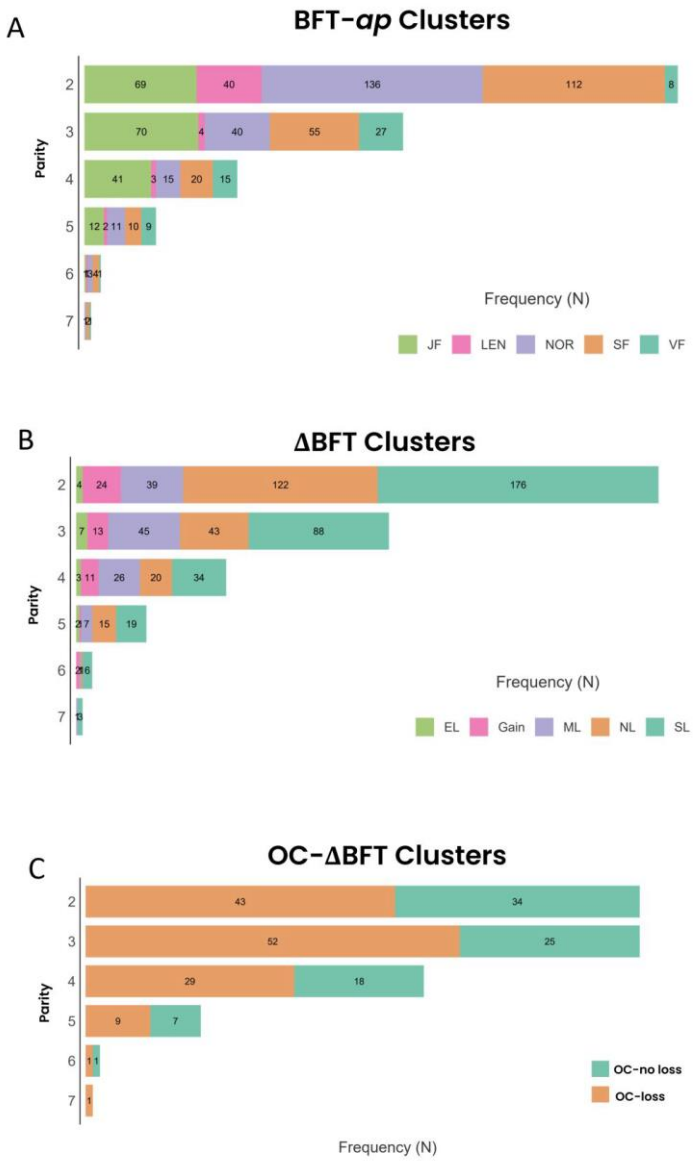
```
} else
{
cat ("xcms set does not contains peaks for all mzML samples.")
}

##mzmatch.ipeak.Combine(sampleList=sampleList, rtwindow=30, combination="set",
ppm=5, outputfolder="combined")


MainClasses <- dir ()
dir.create ("combined_RSD_filtered")
dir.create ("combined_RSD_rejected")
dir.create ("combined")
for (i in 1:length(MainClasses)){FILESf <- dir
(MainClasses[i],full.names=TRUE,pattern="\\.peakml$",recursive=TRUE)
OUTPUTf <- paste ("combined/",MainClasses[i],".peakml",sep="")
if(length(FILESf)>0){mzmatch.ipeak.Combine
(i=paste(FILESf,collapse=","),v=T,rtwindow=30,o=OUTPUTf,combination="set",ppm=5,l
abel=paste(MainClasses[i],sep=""))
RSDf <- paste ("combined_RSD_filtered/",MainClasses[i],".peakml",sep="")
REJf <- paste ("combined_RSD_rejected/",MainClasses[i],".peakml",sep="")
if(length(FILESf)>1)
mzmatch.ipeak.filter.RSDFilter(i=OUTPUTf,o=RSDf,rejected=REJf,rsd=10,v=T) else
file.copy(OUTPUTf,RSDf)}}
INPUTDIR <- "combined_RSD_filtered"
FILESf <- dir (INPUTDIR,full.names=TRUE,pattern="\\.peakml$")
mzmatch.ipeak.Combine(i=paste(FILESf,collapse=","),v=T,rtwindow=30,o="combined.pe
akml",combination="set",ppm=5)

mzmatch.ipeak.filter.SimpleFilter(i="combined.peakml",
o="filteredMinDetectionsPooled.peakml", mindetections=30, setnames="pooled")

mzmatch.ipeak.filter.NoiseFilter(i="combined.peakml",o="combined_noisef.peakml",coda
dw=0.8,)

mzmatch.ipeak.filter.SimpleFilter(i="combined_noisef.peakml",
o="combined_highintensity.peakml", minintensity=100000)

##basepeak extract for file size reduction
mzmatch.ipeak.sort.RelatedPeaks (i="combined_highintensity.peakml", v=T,
o="mzMatch_output.peakml",basepeaks="mzMatch_basepeaks.peakml",ppm=5,rtwindow
=6)

##Gapfill
PeakML.GapFiller(filename = "combined_highintensity.peakml", ionisation = "detect",
Rawpath = NULL, outputfile = "combined_highintensity_gapfilled.peakml", ppm = 0,
rtwin = 0, fillAll=TRUE)

##BatchCorrect
batchCorrect(groups, files, inputFile, outputFile, sampleTypesToKeep, qcLabel="pooled")
```

```
mzmatch.ipeak.sort.IdentifyPeaksets(i="combined_highintensity_gapfilled_corrected.peak
ml", o="Corrected_Identified.peakml")

annot <-
paste("relation.id,relation.ship,codadw,charge,id,identification,adduct,moleculeName")
DBS <- dir(paste(find.package("mzmatch.R"), "/dbs", sep=""), full.names=TRUE)
DBS
DBS <- paste(DBS[c(4,7,8,9)],collapse=",")

mzmatch.ipeak.util.Identify(i="Corrected_Identified.peakml",
o="Corrected_Identified_DBS.peakml", ppm=5, databases=DBS)

##create text file for uncorrected data
mzmatch.ipeak.convert.ConvertToText (i="mzMatch_basepeaks.peakml",
o="uncorrectedOutput.txt",v=T,annotations=annot)

##create text file for corrected data
mzmatch.ipeak.convert.ConvertToText (i="Corrected_Identified_DBS.peakml",
o="Corrected_Identified_DBS.txt",annotations=annot)


##Added text converts if required for trouble-shooting

##mzmatch.ipeak.convert.ConvertToText (i="combined_highintensity.peakml",
o="combined_highintensity.txt",annotations=annot)
##mzmatch.ipeak.convert.ConvertToText (i="combined_highintensity_gapfilled.peakml",
o="combined_highintensity_gapfilled.txt",annotations=annot)
##mzmatch.ipeak.convert.ConvertToText (i="combined.peakml",
o="combined.txt",annotations=annot)
```

A.2 Figures



Supplemental Figure S1. *Frequency (N) of cows supplemented with or without rumen-protected methionine in different clusters. Five BFT-ap clusters: [LEN: lean, BFT 5 to 8 mm; NOR: normal, BFT 9 to 12 mm; SF: slightly fat, BFT 13 to 16 mm; JF: just fat, 16 to 22 mm; VF: very fat, 23 to 43 mm]. Five ΔBFT clusters [EL: extreme loss (ΔBFT from 17 to 23 mm; ML: moderate loss (9 to 15 mm); SL: small loss (4 to 8 mm); NL: no loss (0 to 3 mm); GN: gain (-8 to -1 mm)], and two OC-ΔBFT clusters: [OC-no-loss = overconditioned cows with little or no BFT loss, OC-loss = overconditioned cows with severe to normal loss].*



Supplemental Figure S2. *Frequency of cows with different parities (2 to 7) in the different clusters. Five BFT-ap clusters: [LEN: lean, BFT 5 to 8 mm; NOR: normal, BFT 9 to 12 mm; SF: slightly fat, BFT 13 to 16 mm; JF: just fat, 16 to 22 mm; VF: very fat, 23 to 43 mm]. Five ΔBFT clusters [EL: extreme loss (ΔBFT from 17 to 23 mm; ML: moderate loss (9 to 15 mm); SL: small loss (4 to 8 mm); NL: no loss (0 to 3 mm); GN: gain (-8 to -1 mm)], and two OC-ΔBFT clusters: [OC-no-loss = overconditioned cows with little or no BFT loss, OC-loss = overconditioned cows with severe to normal loss].*

A.3 Tables

Supplemental Table S1. *Description of the distribution of the Methionine treatment, endometritis, and hyperketonaemia as described by Süss et al. 2019, following our clustering based on the BFT-ap.*

Item	Very fat	Just fat	Slightly fat	Normal	Lean
BFT (mm)	23-43	17-22	13-16	9 - 12	5 - 8
Number of cows	61	193	203	206	50
Diagnoses (5 DIM)					
Puerperal metritis[1], (%)	6 (9.8)	6 (3.1)	7 (3.4)	4 (1.9)	6 (12.0)
Clinical metritis[2], n (%)	2 (3.3)	25 (13.0)	33 (16.3)	40 (19.4)	13 (26.0)
Diagnoses (31 DIM)					
Subclinical endometritis[3], n (%)	11 (18.0)	33 (17.1)	32 (15.8)	34 (16.5)	9 (18.0)
Clinical endometritis[4], n (%)	14 (23.0)	48 (24.9)	53 (26.1)	49 (23.8)	12 (24.0)
Hyperketonaemic [5] (3, 5, and 8 DIM)	12 (19.7)	25 (13.0)	33 (16.3)	40 (19.4)	13 (26.0)

[1] Cows were classified based on vaginal discharge and body temperature > 39.5 °C as puerperal metritis
[2] Cows were classified based on vaginal discharge score and body temperature ≤ 39.5 °C as puerperal metritis
[3]Cows were classified based on a vaginal discharge score and uterine cytology. Cows with clear mucus and ≥ 5% polymorhonuclear neutrophils (PMN) in the cytobrush sample were classified as being affected with subclinical endometritis
[4] Cows with vaginal discharge containing off-white or white mucopurulent material and ≥ 5% polymorhonuclear neutrophils (PMN) in the cytobrush sample were classified as being affected with clinical endometritis
[5] Hyperketonaemic: Serum concentrations of BHB on DIM 3, 5 and 8 at least once ≥ 1.2 mM
[1 – 5] the classification with the underlying assessments are described in detail by Süss et al. (2019)

Supplemental Table S2. *Calving to conception interval (CI) calculated for the BFT-ap-clustered animals. CI-pre represents the CI interval of the cows before our trial, whilst the CI after represents the CI subsequent to our study.*

Item	Cluster (BFT-ap)[1]						P-value			
	VF	JF	SF	NOR	LEN	SEM	Cluster	MET	Uterine health[2]	Cluster × MET
CI current (d)	383	384	378	371	372	7,8	0.35	0.27	0.83	0.18
CI after (d)	391	389	387	385	378	9,6	0.85	0.34	0.34	0.18

[1] Cluster: VF= very fat, 23 to 43 mm; JF = just fat, JF, 16 to 22 mm; SF = slightly fat, BFT 13 to 16 mm; NOR= normal, BFT 9 to 12 mm; LEN = lean, BFT 5 to 8 mm.
[2]Uterine health = cows were classified as healthy or affected by endometritis based on vaginal examination and uterine cytology. Vaginal discharge score = clear mucus proportion of PMN < 5% or affected by endometritis: subclinical endometritis (vaginal discharge score = ≤ 50% off-white or white, proportion of PMN ≥ 5%) or clinical endometritis (≥ 50% off-white or white mucopurulent material). PMN = proportion of polymorphonuclear neutrophils.

Supplemental Table S3. *Description of the distribution of the Methionine treatment, endometritis, and hyperketonemia as described by Süss et al. 2019, following our clustering based on the ΔBFT.*

Item	**Extreme Loss**	**Moderate Loss**	**Small Loss**	**No Loss**	**Gain**
ΔBFT (mm)	17 - 23	9 - 15	4 - 8	0 - 3	-8 - -1
Number of cows	16	119	326	201	51
Diagnoses (5 DIM)					
Puerperal metritis[1], n (%)	2 (12.5)	9 (7.6)	7 (2.1)	10 (5.0)	1 (2.0)
Clinical metritis[2], n (%)	2 (12.5)	13 (10.9)	48 (14.7)	38 (18.9)	12 (23.5)
Diagnoses (31 DIM)					
Subclinical endometritis[3], n (%)	3 (18.8)	16 (13.4)	59 (18.1)	31 (15.4)	10 (19.6)
Clinical endometritis[4], n (%)	5 (31.3)	31 (26.1)	82 (25.2)	49 (24.4)	9 (17.6)
Hyperketonaemic [5] (3, 5, and 8 DIM)	3 (18.8)	24 (20.2)	36 (11.0)	9 (4.5)	1 (2.0)

[1] The cows were classified based on vaginal discharge and body temperature > 39.5 °C as puerperal metritis
[2] The cows were classified based on vaginal discharge score and body temperature ≤ 39.5 °C as puerperal metritis
[3]The cows were classified based on a vaginal discharge score and uterine cytology. Cows with clear mucus and ≥5% polymorhonuclear neutrophils (PMN) in the cytobrush sample were classified as being affected with subclinical endometritis
[4] Cows with vaginal discharge containing off-white or white mucopurulent material and ≥5% polymorhonuclear neutrophils (PMN) in the cytobrush sample were classified as being affected with clinical endometritis
[5] Hyperketonaemic: Serum concentrations of BHB on DIM 3, 5, and 8 at least once ≥ 1.2 mM
[1 – 5] the classification with the underlying assessments was described in detail by Süss et al. (2019)

Supplemental Table S4. List of the matched metabolites for the Metabolomics analysis performed using PiMP. The analysis on the BC loss led to the identification of 76 metabolites for the Batch 1 and 82 metabolites for the Batch 2, whilst the analysis of the Control vs Methionine supplemented diet led to the identification of 65 metabolites.

Peak ID	Name	Formula
BC Loss Batch 1		
9	Betaine	C5H11NO2
10	Creatinine	C4H7N3O
21	L-Proline	C5H9NO2
24	L-Glutamine	C5H10N2O3
46	L-Leucine	C6H13NO2
51	L-Valine	C5H11NO2
58	L-Arginine	C6H14N4O2
63	L-isoleucine	C6H13NO2
65	Nicotinamide	C6H6N2O
66	L-Citrulline	C6H13N3O3
75	L-Phenylalanine	C9H11NO2
80	N(pi)-Methyl-L-histidine	C7H11N3O2
81	L-Carnitine	C7H15NO3
84	O-Acetylcarnitine	C9H17NO4
90	Choline	C5H13NO
103	beta-Alanine	C3H7NO2
103	L-Alanine	C3H7NO2
111	L-Methionine	C5H11NO2S
147	L-homoserine	C4H9NO3
147	L-Threonine	C4H9NO3
149	L-Tryptophan	C11H12N2O2
158	trans-4-Hydroxy-L-proline	C5H9NO3
158	5-Aminolevulinate	C5H9NO3
162	L-Tyrosine	C9H11NO3
185	L-Glutamate	C5H9NO4
185	O-Acetyl-L-serine	C5H9NO4
193	L-Ornithine	C5H12N2O2
216	L-Lysine	C6H14N2O2
222	Phenylacetylglycine	C10H11NO3
229	Glycine	C2H5NO2
292	N-Acetyl-D-glucosamine	C8H15NO6

299	Isonicotinic acid	C6H5NO2
299	Nicotinate	C6H5NO2
300	Cytidine	C9H13N3O5
303	sn-glycero-3-Phosphocholine	C8H20NO6P
356	Pantothenate	C9H17NO5
368	L-Asparagine	C4H8N2O3
690	D-Glucosamine	C6H13NO5
850	N-Acetylglutamine	C7H12N2O4
881	Adenine	C5H5N5
1133	(R)-Lactate	C3H6O3
1142	(R)-3-Hydroxybutanoate	C4H8O3
1143	citrate	C6H8O7
1189	5-Oxoproline	C5H7NO3
1254	(S)-Malate	C4H6O5
1290	Taurine	C2H7NO3S
1325	L-Serine	C3H7NO3
1499	allantoin	C4H6N4O3
1503	sucrose	C12H22O11
1503	Maltose	C12H22O11
1572	D-glucose	C6H12O6
1572	D-Fructose	C6H12O6
1588	Uridine	C9H12N2O6
1616	(R)-2-Hydroxyglutarate	C5H8O5
1636	D-Gluconic acid	C6H12O7
1662	Deoxyuridine	C9H12N2O5
1698	Phthalate	C8H6O4
1702	Pyruvate	C3H4O3
1706	thymine	C5H6N2O2
1727	Oxalate	C2H2O4
1789	Methylmalonate	C4H6O4
1789	Succinate	C4H6O4
1868	D-Galacturonate	C6H10O7
1890	Maleic acid	C4H4O4
1932	D-Erythrose	C4H8O4
1953	cis-Aconitate	C6H6O6

1983	L-Aspartate	C4H7NO4
2046	Orotate	C5H4N2O4
2062	2-Methylcitrate	C7H10O7
2103	Itaconate	C5H6O4
2170	L-Kynurenine	C10H12N2O3
2185	Malonate	C3H4O4
2322	L-Cystine	C6H12N2O4S2
2335	L-2-Aminoadipate	C6H11NO4
2353	Methylcysteine	C4H9NO2S
2458	L-Cystathionine	C7H14N2O4S
BC Loss Batch 2		
10	Betaine	C5H11NO2
11	Creatinine	C4H7N3O
22	L-Glutamine	C5H10N2O3
25	L-Proline	C5H9NO2
29	L-Leucine	C6H13NO2
38	L-Valine	C5H11NO2
43	L-Arginine	C6H14N4O2
54	L-Citrulline	C6H13N3O3
63	L-isoleucine	C6H13NO2
66	L-Carnitine	C7H15NO3
69	O-Acetylcarnitine	C9H17NO4
70	L-Phenylalanine	C9H11NO2
78	N(pi)-Methyl-L-histidine	C7H11N3O2
84	beta-Alanine	C3H7NO2
84	L-Alanine	C3H7NO2
87	Choline	C5H13NO
92	L-Methionine	C5H11NO2S
120	Phenylacetylglycine	C10H11NO3
121	L-Tryptophan	C11H12N2O2
127	L-Tyrosine	C9H11NO3
128	trans-4-Hydroxy-L-proline	C5H9NO3
128	5-Aminolevulinate	C5H9NO3
153	L-Lysine	C6H14N2O2
169	L-Ornithine	C5H12N2O2

181	4-Trimethylammoniobutanoate	C7H15NO2
182	Hypoxanthine	C5H4N4O
184	Inosine	C10H12N4O5
189	L-Glutamate	C5H9NO4
189	O-Acetyl-L-serine	C5H9NO4
204	L-Kynurenine	C10H12N2O3
214	Cytidine	C9H13N3O5
227	L-Asparagine	C4H8N2O3
227	Glycylglycine	C4H8N2O3
235	sn-glycero-3-Phosphocholine	C8H20NO6P
238	Glycine	C2H5NO2
370	Pyridoxal	C8H9NO3
505	cytosine	C4H5N3O
542	Pantothenate	C9H17NO5
623	Isonicotinic acid	C6H5NO2
623	Nicotinate	C6H5NO2
625	Imidazole-4-acetate	C5H6N2O2
967	L-Cystine	C6H12N2O4S2
993	N-Acetylglutamine	C7H12N2O4
1296	L-Cystathionine	C7H14N2O4S
1435	(R)-Lactate	C3H6O3
1437	(R)-3-Hydroxybutanoate	C4H8O3
1468	citrate	C6H8O7
1525	5-Oxoproline	C5H7NO3
1542	Taurine	C2H7NO3S
1549	(S)-Malate	C4H6O5
1550	Taurocholate	C26H45NO7S
1565	L-homoserine	C4H9NO3
1565	L-Threonine	C4H9NO3
1570	allantoin	C4H6N4O3
1597	(R)-2-Hydroxyglutarate	C5H8O5
1599	L-Serine	C3H7NO3
1670	D-glucose	C6H12O6
1806	D-Fructose	C6H12O6
1858	Uridine	C9H12N2O6

1881	D-Gluconic acid	C6H12O7
1885	Orotate	C5H4N2O4
1910	thymine	C5H6N2O2
1917	Orthophosphate	H3O4P
1928	Pyruvate	C3H4O3
1933	2-oxobutanoate	C4H6O3
1935	Deoxyuridine	C9H12N2O5
1951	Methylmalonate	C4H6O4
1951	Succinate	C4H6O4
2000	D-Galacturonate	C6H10O7
2008	Oxalate	C2H2O4
2014	Maleic acid	C4H4O4
2016	L-Aspartate	C4H7NO4
2040	L-Rhamnose	C6H12O5
2280	D-Erythrose	C4H8O4
2332	D-Threose	C4H8O4
2369	Malonate	C3H4O4
2411	2-Methylcitrate	C7H10O7
2514	L-2-Aminoadipate	C6H11NO4
2646	Methylcysteine	C4H9NO2S
2714	N-acetyl-L-glutamate	C7H11NO5
2751	D-glucose 6-phosphate	C6H13O9P
2751	D-Fructose 6-phosphate	C6H13O9P
Diet		
8	Betaine	C5H11NO2
9	Creatinine	C4H7N3O
19	L-Proline	C5H9NO2
22	L-Glutamine	C5H10N2O3
41	L-Leucine	C6H13NO2
46	L-Valine	C5H11NO2
52	L-Arginine	C6H14N4O2
55	L-isoleucine	C6H13NO2
56	L-Citrulline	C6H13N3O3
62	L-Phenylalanine	C9H11NO2
66	N(pi)-Methyl-L-histidine	C7H11N3O2

67	L-Carnitine	C7H15NO3
69	O-Acetylcarnitine	C9H17NO4
74	Choline	C5H13NO
78	beta-Alanine	C3H7NO2
78	L-Alanine	C3H7NO2
79	L-Methionine	C5H11NO2S
99	L-homoserine	C4H9NO3
99	L-Threonine	C4H9NO3
101	L-Tryptophan	C11H12N2O2
108	trans-4-Hydroxy-L-proline	C5H9NO3
108	5-Aminolevulinate	C5H9NO3
112	L-Tyrosine	C9H11NO3
125	L-Kynurenine	C10H12N2O3
126	L-Glutamate	C5H9NO4
126	O-Acetyl-L-serine	C5H9NO4
134	L-Ornithine	C5H12N2O2
150	L-Lysine	C6H14N2O2
160	Glycine	C2H5NO2
216	Cytidine	C9H13N3O5
218	sn-glycero-3-Phosphocholine	C8H20NO6P
273	Pantothenate	C9H17NO5
367	Imidazole-4-acetate	C5H6N2O2
495	D-Glucosamine	C6H13NO5
607	N-Acetylglutamine	C7H12N2O4
629	Adenine	C5H5N5
798	(R)-Lactate	C3H6O3
804	(R)-3-Hydroxybutanoate	C4H8O3
805	citrate	C6H8O7
822	5-Oxoproline	C5H7NO3
889	(S)-Malate	C4H6O5
891	Taurine	C2H7NO3S
907	L-Serine	C3H7NO3
993	allantoin	C4H6N4O3
1026	2-oxobutanoate	C4H6O3
1064	Uridine	C9H12N2O6

1100	(R)-2-Hydroxyglutarate	C5H8O5
1157	D-Gluconic acid	C6H12O7
1196	Deoxyuridine	C9H12N2O5
1274	Pyruvate	C3H4O3
1276	thymine	C5H6N2O2
1287	Oxalate	C2H2O4
1330	Methylmalonate	C4H6O4
1330	Succinate	C4H6O4
1349	L-Asparagine	C4H8N2O3
1386	D-Galacturonate	C6H10O7
1391	Maleic acid	C4H4O4
1408	D-Erythrose	C4H8O4
1455	L-Aspartate	C4H7NO4
1500	Orotate	C5H4N2O4
1540	Itaconate	C5H6O4
1594	Malonate	C3H4O4
1694	L-Cystine	C6H12N2O4S2
1702	L-2-Aminoadipate	C6H11NO4
1720	Methylcysteine	C4H9NO2S

Supplemental Table S5. *List of serum proteins, ranked from most significant to least significant difference, when comparing fat (> 17 mm BFT-ap) and lean (< 13 mm BFT-ap) cows around 30 days pp before batch correction.*

Accession number	Protein name	logFC	Adjusted P-values	Number of peptides
A5PJ69;ENSEMBL:ENSBTAP00000013050	Serpina-10	-0,48837	0,027493064	19
A0A3Q1MQV3;F1N3A1;Q28178;Q28194	Thrombospondin-1	0,612369	0,035821363	33
A0A140T897	Albumin	0,329112	0,041883647	79
P80195	Glycam-1	-0,42626	0,041883647	8
Q0IIK2	Serotransferrin	0,282352	0,108636396	62
A0A3Q1NJR8	Antithrombin-III	-0,41649	0,108636396	42
A0A3Q1LPG0	Uncharacterized	0,415912	0,152854686	16
P28800	Alpha-2-antiplasmin	-0,25557	0,161457242	28
O46375	Transthyretin	0,530579	0,255958023	11
A0A452DIP8;P00743	Coagulation factor X	-0,25321	0,258126154	7
Q95M17	Acidic mammalian chitinase	-0,45715	0,285758818	6
E1B805	Uncharacterized	-0,23359	0,285758818	46
Q7SIH1	Alpha-2-macroglobulin	0,153953	0,285758818	120
G5E5V1	Ig-like domain-containing protein	0,341764	0,285758818	2
P00735	Prothrombin	0,214239	0,285758818	18
A0A3Q1LSR2;A0A3Q1MWQ1;Q2TBQ1	Coagulation factor XIII B chain	0,184355	0,288404294	7
ENSEMBL:ENSBTAP00000024466	Uncharacterized	0,25933	0,288404294	24
F1MU18	Oncostatin M receptor	-0,23532	0,288404294	5
A0A452DJK6;Q9TTE1	Serpin A3-1	0,214363	0,288404294	14
A0A3Q1LJT1	Ig-like domain-containing protein	0,379165	0,288404294	4

Supplemental Table S6. *List of 20 top proteins, ranked from most significant to least significant, when comparing fat (> 17 mm BFT-ap) and lean (< 13 mm BFT-ap) cows around 30 days pp, after batch correction.*

Accession number	Protein name	logFC	Adjusted P-values	Number of peptides
O02659	Mannose-binding protein C	-0,332704791	0,238923619	5
A5PJ69;ENSEMBL:ENSBTAP00000013050	Serpina-10	-0,408026729	0,238923619	19
A0A140T897	Albumin	0,310556949	0,238923619	79
A0A3Q1MQV3;F1N3A1;Q28178;Q28194	Thrombospondin-1	0,619818428	0,238923619	33
P80195	Glycam-1	-0,416668089	0,242044758	8
Q0IIK2	Serotransferrin	0,262609702	0,3692708	62
A0A3Q1LSR2;A0A3Q1MWQ1;Q2TBQ1	Coagulation factor XIII B chain	0,232170446	0,3692708	7
A0A452DHX8	Amine oxidase	0,34725141	0,3692708	9
E1BF81;ENSEMBL:ENSBTAP00000023402	*Uncharacterized*	-0,386020093	0,3692708	13
A0A3Q1NJR8	Antithrombin-III	-0,359122724	0,389884694	42
A0A452DHZ7;Q3T0E5	Adipocyte plasma membrane-associated protein	0,312742937	0,389884694	6
P80012	von Willebrand factor	0,397288548	0,389884694	5
G5E5V1	Ig-like domain-containing protein	0,375108439	0,389884694	2
A8E654;F1N6W9	COL18A1 protein	0,267779017	0,423381125	2
P00735	Prothrombin	0,202152344	0,538574188	18
O97764	Zeta-crystallin	1,082213139	0,580251316	2
A0A3Q1LPG0	*Uncharacterized*	0,328847396	0,580251316	16
F1N1I6;Q3SX14	Actin-depolymerizing factor	0,158086124	0,580251316	48
A0A452DIP8;P00743	Coagulation factor X	-0,210069018	0,580251316	7
E1BFN6	Dihydropyrimidinase	-0,418413614	0,580251316	2

References

Abuajamieh, M., S.K. Kvidera, M.V.S. Fernandez, A. Nayeri, N.C. Upah, E.A. Nolan, S.M. Lei, J.M. DeFrain, H.B. Green, K.M. Schoenberg, W.E. Trout, and L.H. Baumgard. 2016. Inflammatory biomarkers are associated with ketosis in periparturient Holstein cows. Res. Vet. Sci. 109:81–85. doi:10.1016/j.rvsc.2016.09.015.

Ahima, R.S., and M.A. Lazar. 2008. Adipokines and the peripheral and neural control of energy balance. Mol. Endocrinol. 22:1023–1031. doi:10.1210/me.2007-0529.

Alharthi, A., Z. Zhou, V. Lopreiato, E. Trevisi, and J.J. Loor. 2018. Body condition score prior to parturition is associated with plasma and adipose tissue biomarkers of lipid metabolism and inflammation in Holstein cows. J. Anim. Sci. Biotechnol. 9:1–12. doi:10.1186/s40104-017-0221-1.

Almeida, A.M., S.A. Ali, F. Ceciliani, P.D. Eckersall, L.E. Hernández-Castellano, R. Han, J.J. Hodnik, S. Jaswal, J.D. Lippolis, M. McLaughlin, I. Miller, A.K. Mohanty, V. Mrljak, J.E. Nally, P. Nanni, J.E. Plowman, M.D. Poleti, D.M. Ribeiro, P. Rodrigues, B. Roschitzki, R. Schlapbach, J. Starič, Y. Yang, and M. Zachut. 2021. Domestic animal proteomics in the 21st century: A global retrospective and viewpoint analysis. J. Proteomics. 241:104220. doi:10.1016/j.jprot.2021.104220.

Ametaj, B.N., B.J. Bradford, G. Bobe, R.A. Nafikov, Y. Lu, J.W. Young, and D.C. Beitz. 2005. Strong relationships between mediators of the acute phase response and fatty liver in dairy cows. Can. J. Anim. Sci. 85:165–175. doi:10.4141/A04-043.

Amon, S., F. Meier-Abt, L.C. Gillet, S. Dimitrieva, A.P.A. Theocharides, M.G. Manz, and R. Aebersold. 2019. Sensitive quantitative proteomics of human hematopoietic stem and progenitor cells by data-independent acquisition mass spectrometry. Mol. Cell. Proteomics 18:1454–1467. doi:10.1074/mcp.TIR119.001431.

Ardalan, M., and K. Rezayazdi. 2010. Effect of rumen-protected choline and methionine on physiological and metabolic disorders and reproductive indices of dairy cows. J. Anim. Physiol. Anim. Nutr. 94:259–265. doi:10.1111/j.1439-0396.2009.00966.x.

Asl, A.N., S. Nazifi, A.R. Ghasrodashti, and A. Olyaee. 2011. Prevalence of subclinical ketosis in dairy cattle in the Southwestern Iran and detection of cutoff point for NEFA and glucose concentrations for diagnosis of subclinical ketosis. Prev. Vet. Med. 100:38–43. doi:10.1016/j.prevetmed.2011.02.013.

Bagheri, M., F. Farzadfar, L. Qi, M.S. Yekaninejad, M. Chamari, O.A. Zeleznik, Z. Kalantar, Z. Ebrahimi, A. Sheidaie, B. Koletzko, O. Uhl, and A. Djazayery. 2018. Obesity-related metabolomic profiles and discrimination of metabolically unhealthy obesity. J. Proteome Res. 17:1452–1462. doi:10.1021/acs.jproteome.7b00802.

Barletta, R. V., M. Maturana Filho, P.D. Carvalho, T.A. Del Valle, A.S. Netto, F.P. Rennó, R.D. Mingoti, J.R. Gandra, G.B. Mourão, P.M. Fricke, R. Sartori, E.H. Madureira, and M.C. Wiltbank. 2017. Association of changes among body condition score during the transition period with NEFA and BHBA concentrations, milk production, fertility, and health of Holstein cows. Theriogenology 104:30–36. doi:10.1016/j.theriogenology.2017.07.030.

Barragan, A.A., J.M. Piñeiro, G.M. Schuenemann, P.J. Rajala-Schultz, D.E. Sanders, J. Lakritz, and S. Bas. 2018. Assessment of daily activity patterns and biomarkers of pain, inflammation, and stress in lactating dairy cows diagnosed with clinical metritis. J. Dairy Sci. 101:8248–8258. doi:10.3168/jds.2018-14510.

Batistel, F., J.M. Arroyo, A. Bellingeri, L. Wang, B. Saremi, C. Parys, E. Trevisi, F.C. Cardoso, and J.J. Loor. 2017. Ethyl-cellulose rumen-protected methionine enhances performance during the periparturient period and early lactation in Holstein dairy cows. J. Dairy Sci. 100:7455–7467. doi:10.3168/jds.2017-12689.

Bauman, D.E., K.J. Harvatine, and A.L. Lock. 2011. Nutrigenomics, rumen-derived bioactive fatty acids, and the regulation of milk fat synthesis. Annu. Rev. Nutr. 31:299–319. doi:10.1146/annurev.nutr.012809.104648.

Baumgard, L.H., J.K. Sangster, and D.E. Bauman. 2001. Milk fat synthesis in dairy cows is progressively reduced by increasing supplemental amounts of trans-10, cis-12 conjugated linoleic acid (CLA). J. Nutr. 131:1764–1769. doi:10.1093/jn/131.6.1764.

Beauchemin, K.A., L. Eriksen, P. Nørgaard, and L.M. Rode. 2008. Short communication: Salivary secretion during meals in lactating dairy cattle. J. Dairy Sci. 91:2077–2081. doi:10.3168/jds.2007-0726.

Bell, A.W. 1995. Regulation of organic nutrient metabolism during transition from late pregnancy to early lactation.. J. Anim. Sci. 73:2804–2819. doi:10.2527/1995.7392804x.

Bell, A.W., W.S. Burhans, and T.R. Overton. 2000. Protein nutrition in late pregnancy, maternal protein reserves and lactation performance in dairy cows. Proc. Nutr. Soc.

59:119–126. doi:10.1017/S0029665100000148.

Bellei, E., S. Bergamini, E. Monari, L.I. Fantoni, A. Cuoghi, T. Ozben, and A. Tomasi. 2011. High-abundance proteins depletion for serum proteomic analysis: Concomitant removal of non-targeted proteins. Amino Acids 40:145–156. doi:10.1007/s00726-010-0628-x.

Bendixen, E., M. Danielsen, K. Hollung, E. Gianazza, and I. Miller. 2011. Farm animal proteomics - A review. J. Proteomics 74:282–293. doi:10.1016/j.jprot.2010.11.005.

Benedet, A., C.L. Manuelian, A. Zidi, M. Penasa, and M. De Marchi. 2019. Invited review: β-hydroxybutyrate concentration in blood and milk and its associations with cow performance. Animal 13:1676–1689. doi:10.1017/S175173111900034X.

Van Den Bergh, G., and L. Arckens. 2005. Recent advances in 2D electrophoresis: An array of possibilities. Expert Rev. Proteomics 2:243–252. doi:10.1586/14789450.2.2.243.

Bernabucci, U., B. Ronchi, N. Lacetera, and A. Nardone. 2005. Influence of body condition score on relationships between metabolic status and oxidative stress in periparturient dairy cows. J. Dairy Sci. 88:2017–2026. doi:10.3168/jds.s0022-0302(05)72878-2.

Bernal-Santos, G., J.W. Perfield, D.M. Barbano, D.E. Bauman, and T.R. Overton. 2003. Production responses of dairy cows to dietary supplementation with Conjugated Linoleic Acid (CLA) during the transition period and early lactation. J. Dairy Sci. 86:3218–3228. doi:10.3168/jds.S0022-0302(03)73925-3.

Bevilacqua, M., and R. Bro. 2020. Can we trust score plots?. Metabolites 10:1–7. doi:10.3390/metabo10070278.

Bilić, P., J. Kuleš, A. Galan, L. Gomes de Pontes, N. Guillemin, A. Horvatić, A. Festa Sabes, V. Mrljak, and P.D. Eckersall. 2018. Proteomics in veterinary medicine and animal science: Neglected scientific opportunities with immediate impact. Proteomics 18:1–7. doi:10.1002/pmic.201800047.

Block, S.S., R.P. Rhoads, D.E. Bauman, R.A. Ehrhardt, M.A. McGuire, B.A. Crooker, J.M. Griinari, T.R. Mackle, W.J. Weber, M.E. Van Amburgh, and Y.R. Boisclair. 2003. Demonstration of a role for insulin in the regulation of leptin in lactating dairy cows. J. Dairy Sci. 86:3508–3515. doi:10.3168/jds.S0022-0302(03)73955-1.

Brunius, C., L. Shi, and R. Landberg. 2016. Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. Metabolomics 12:1–13. doi:10.1007/s11306-016-1124-4.

Cao, G., Z. Song, Y. Hong, Z. Yang, Y. Song, Z. Chen, Z. Chen, and Z. Cai. 2020. Large-scale targeted metabolomics method for metabolite profiling of human samples. Anal. Chim. Acta 1125:144–151. doi:10.1016/j.aca.2020.05.053.

Capuco, A. V., R.M. Akers, and J.J. Smith. 1997. Mammary growth in holstein cows during the dry period: Quantification of nucleic acids and histology. J. Dairy Sci. 80:477–487. doi:10.3168/jds.S0022-0302(97)75960-5.

Cardoso, F.F., S.S. Donkin, M.N. Pereira, R.A.N. Pereira, A.P. Peconick, J.P. Santos, R.B. Silva, V.R. Caproni, C. Parys, and M.A.C. Danes. 2021. Effect of protein level and methionine supplementation on dairy cows during the transition period. J. Dairy Sci. 104:5467–5478. doi:10.3168/jds.2020-19181.

Carlson, D.B., N.B. Litherland, H.M. Dann, J.C. Woodworth, and J.K. Drackley. 2006. Metabolic effects of abomasal L-carnitine infusion and feed restriction in lactating holstein cows. J. Dairy Sci. 89:4819–4834. doi:10.3168/jds.S0022-0302(06)72531-0.

Carlson, D.B., J.W. McFadden, A. D'Angelo, J.C. Woodworth, and J.K. Drackley. 2007. Dietary L-carnitine affects periparturient nutrient metabolism and lactation in multiparous cows. J. Dairy Sci. 90:3422–3441. doi:10.3168/jds.2006-811.

Carvalho, P.D., A.H. Souza, M.C. Amundson, K.S. Hackbart, M.J. Fuenzalida, M.M. Herlihy, H. Ayres, A.R. Dresch, L.M. Vieira, J.N. Guenther, R.R. Grummer, P.M. Fricke, R.D. Shaver, and M.C. Wiltbank. 2014. Relationships between fertility and postpartum changes in body condition and body weight in lactating dairy cows. J. Dairy Sci. 97:3666–3683. doi:10.3168/jds.2013-7809.

Castañeda-Gutiérrez, E., B.C. Benefield, M.J. De Veth, N.R. Santos, R.O. Gilbert, W.R. Butler, and D.E. Bauman. 2007. Evaluation of the mechanism of action of conjugated linoleic acid isomers on reproduction in dairy cows. J. Dairy Sci. 90:4253–4264. doi:10.3168/jds.2007-0117.

Ceciliani, F., J.J. Ceron, P.D. Eckersall, and H. Sauerwein. 2012. Acute phase proteins in ruminants. J. Proteomics 75:4207–4231. doi:10.1016/j.jprot.2012.04.004.

Ceciliani, F., C. Lecchi, C. Urh, and H. Sauerwein. 2018a. Proteomics and metabolomics

characterizing the pathophysiology of adaptive reactions to the metabolic challenges during the transition from late pregnancy to early lactation in dairy cows. J. Proteomics 178:92–106. doi:10.1016/j.jprot.2017.10.010.

Ceciliani, F., C. Lecchi, C. Urh, and H. Sauerwein. 2018b. Proteomics and metabolomics characterizing the pathophysiology of adaptive reactions to the metabolic challenges during the transition from late pregnancy to early lactation in dairy cows. J. Proteomics 178:92–106. doi:10.1016/j.jprot.2017.10.010.

Chait, B.T. 2011. Mass spectrometry in the postgenomic Era. Annu. Rev. Biochem. 80:239–246. doi:10.1146/annurev-biochem-110810-095744.

Chandler, T.L., and H.M. White. 2017. Choline and methionine differentially alter methyl carbon metabolism in bovine neonatal hepatocytes. PLoS One 12:1–16. doi:10.1371/journal.pone.0171080.

Chandramouli, K., and P.-Y. Qian. 2009. Proteomics: Challenges, techniques and possibilities to overcome biological sample complexity. Hum. Genomics Proteomics 1. doi:10.4061/2009/239204.

Chen, Z.H., G.A. Broderick, N.D. Luchini, B.K. Sloan, and E. Devillard. 2011. Effect of feeding different sources of rumen-protected methionine on milk production and N-utilization in lactating dairy cows. J. Dairy Sci. 94:1978–1988. doi:10.3168/jds.2010-3578.

Chilliard, Y., C. Delavaud, and M. Bonnet. 2005. Leptin expression in ruminants: Nutritional and physiological regulations in relation with energy metabolism. Domest. Anim. Endocrinol. 29:3–22. doi:10.1016/j.domaniend.2005.02.026.

Choe, C., J.W. Park, E.S. Kim, S.G. Lee, S.Y. Park, J.S. Lee, M.J. Cho, K.R. Kang, J. Han, and D. Kang. 2010. Proteomic analysis of differentially expressed proteins in bovine endometrium with endometritis. Korean J. Physiol. Pharmacol. 14:205–212. doi:10.4196/kjpp.2010.14.4.205.

Chong, J., D.S. Wishart, and J. Xia. 2019. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. Curr. Protoc. Bioinforma. 68:1–128. doi:10.1002/cpbi.86.

Clish, C.B. 2015. Metabolomics: an emerging but powerful tool for precision medicine. Mol. Case Stud. 1:a000588. doi:10.1101/mcs.a000588.

Coble, J.B., and C.G. Fraga. 2014. Comparative evaluation of preprocessing freeware on chromatography/mass spectrometry data for signature discovery. J. Chromatogr. A 1358:155–164. doi:10.1016/j.chroma.2014.06.100.

Cooke, R.F., N. Silva Del Río, D.Z. Caraviello, S.J. Bertics, M.H. Ramos, and R.R. Grummer. 2007. Supplemental choline for prevention and alleviation of fatty liver in dairy cattle. J. Dairy Sci. 90:2413–2418. doi:10.3168/jds.2006-028.

Council, N.R. 2001. Nutrient Requirements of Dairy Cattle: Seventh Revised Edition, 2001. The National Academies Press, Washington, DC.

Daddam, J.R., H.M. Hammon, A. Tröscher, L. Vogel, M. Gnott, G. Kra, Y. Levin, H. Sauerwein, and M. Zachut. 2021. Phosphoproteomic analysis of subcutaneous and omental adipose tissue reveals increased lipid turnover in dairy cows supplemented with conjugated linoleic acid. Int. J. Mol. Sci. 22:1–26. doi:10.3390/ijms22063227.

Das, S., A. Dey, A. Pal, and N. Roy. 2015. Applications of artificial intelligence in machine learning: review and prospect. Int. J. Comput. Appl. 115:31–41. doi:10.5120/20182-2402.

Depreester, E., J. De Koster, M. Van Poucke, M. Hostens, W. Van den Broeck, L. Peelman, G.A. Contreras, and G. Opsomer. 2018. Influence of adipocyte size and adipose depot on the number of adipose tissue macrophages and the expression of adipokines in dairy cows at the end of pregnancy. J. Dairy Sci. 101:6542–6555. doi:10.3168/jds.2017-13777.

Deutsch, E.W. 2010. Mass Spectrometer Output File Format mzML. Methods Mol Biol. 604:319-31. doi: 10.1007/978-1-60761-444-9_22..

Dey, A. 2016. Machine learning algorithms: A review. Int. J. Comput. Sci. Inf. Technol. 7:1174–1179.

Dhall, D., R. Kaur, and M. Juneja. 2020. Machine Learning: A review of the algorithms and its applications. Pages 47–63 in Proceedings of ICRIC 2019. Springer International Publishing, Cham.

Drackley, J.K. 1999. Biology of dairy cows during the transition period: the final frontier?. J. Dairy Sci. 82:2259–2273. doi:10.3168/jds.S0022-0302(99)75474-3.

Drackley, J.K., and F.C. Cardoso. 2014. Prepartum and postpartum nutritional management to optimize fertility in high-yielding dairy cows in confined TMR

systems. Animal 8:5–14. doi:10.1017/S1751731114000731.

Duffield, T.F., K.D. Lissemore, B.W. McBride, and K.E. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92:571–580. doi:10.3168/jds.2008-1507.

Duffield, T.F., D. Sandals, K.E. Leslie, K. Lissemore, B.W. McBride, J.H. Lumsden, P. Dick, and R. Bagg. 1998. Efficacy of Monensin for the prevention of subclinical ketosis in lactating dairy cows. J. Dairy Sci. 81:2866–2873. doi:10.3168/jds.S0022-0302(98)75846-1.

Eckersall, P.D. 1995. Acute phase proteins as markers of inflammatory lesions. Comp. Haematol. Int. 5:93–97. doi:10.1007/BF00638925.

Edmonson, A.J., I.J. Lean, L.D. Weaver, T. Farver, and G. Webster. 1989. A body condition scoring chart for holstein dairy cows. J. Dairy Sci. 72:68–78. doi:10.3168/jds.S0022-0302(89)79081-0.

Emery, R.S. 1971. Disappearance of Methionine from the rumen. J. Dairy Sci. 54:1090–1091. doi:10.3168/jds.S0022-0302(71)85978-7.

Erdman, R.A., R.D. Shaver, and J.H. Vandersall. 1984. Dietary choline for the lactating cow: Possible effects on milk fat synthesis. J. Dairy Sci. 67:410–415. doi:10.3168/jds.S0022-0302(84)81317-X.

Esposito, G., V.A. Absalón Medina, A. Schneider, R.O. Gilbert, and W.R. Butler. 2013. Effect of dietary conjugated linoleic acid (CLA) on the metabolism and reproduction of dairy cows. South African J. Anim. Sci. 43. doi:10.4314/sajas.v43i5.6.

Esposito, G., P.C. Irons, E.C. Webb, and A. Chapwanya. 2014. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. Anim. Reprod. Sci. 144:60–71. doi:10.1016/j.anireprosci.2013.11.007.

Fan, Z., S. Shu, C. Xu, X. Xiao, G. Wang, Y. Bai, C. Xia, L. Wu, H. Zhang, C. Xu, and W. Yang. 2017. Protein profiling of plasma proteins in dairy cows with subclinical hypocalcaemia. Ir. Vet. J. 70:1–7. doi:10.1186/s13620-017-0082-0.

Galic, S., J.S. Oakhill, and G.R. Steinberg. 2010. Adipose tissue as an endocrine organ. Mol. Cell. Endocrinol. 316:129–139. doi:10.1016/j.mce.2009.08.018.

Gärtner, T., E. Gernand, J. Gottschalk, and K. Donat. 2019. Relationships between body

condition, body condition loss, and serum metabolites during the transition period in primiparous and multiparous cows. J. Dairy Sci. 102:9187–9199. doi:10.3168/jds.2018-15762.

Geyer, P.E., N.A. Kulak, G. Pichler, L.M. Holdt, D. Teupser, and M. Mann. 2016. Plasma proteome profiling to assess human health and disease. Cell Syst. 2:185–195. doi:10.1016/j.cels.2016.02.015.

Ghaffari, M.H., A. Jahanbekam, C. Post, H. Sadri, K. Schuh, C. Koch, and H. Sauerwein. 2020a. Discovery of different metabotypes in overconditioned dairy cows by means of machine learning. J. Dairy Sci. 103:9604–9619. doi:10.3168/jds.2020-18661.

Ghaffari, M.H., A. Jahanbekam, H. Sadri, K. Schuh, G. Dusel, C. Prehn, J. Adamski, C. Koch, and H. Sauerwein. 2019. Metabolomics meets machine learning: Longitudinal metabolite profiling in serum of normal versus overconditioned cows and pathway analysis. J. Dairy Sci. 102:11561–11585. doi:10.3168/jds.2019-17114.

Ghaffari, M.H., K. Schuh, J. Kuleš, N. Guillemin, A. Horvatić, V. Mrljak, P.D. Eckersall, G. Dusel, C. Koch, H. Sadri, and H. Sauerwein. 2020b. Plasma proteomic profiling and pathway analysis of normal and overconditioned dairy cows during the transition from late pregnancy to early lactation. J. Dairy Sci. 103:4806–4821. doi:10.3168/jds.2019-17897.

Ghaffari, M. H., M. T. Alaedin, H. Sadri, I. Hofs, C. Koch, and H. Sauerwein. 2021. Longitudinal changes in fatty acid metabolism and in the mitochondrial protein import system in overconditioned and normal conditioned cows: A transcriptional study using microfluidic quantitative PCR. J Dairy Sci. https://doi.org/10.3168/jds.2021-20237

Gidlund, H., M. Hetta, S.J. Krizsan, S. Lemosquet, and P. Huhtanen. 2015. Effects of soybean meal or canola meal on milk production and methane emissions in lactating dairy cows fed grass silage-based diets. J. Dairy Sci. 98:8093–8106. doi:10.3168/jds.2015-9757.

Giesy, S.L., B. Yoon, W.B. Currie, J.W. Kim, and Y.R. Boisclair. 2012. Adiponectin deficit during the precarious glucose economy of early lactation in dairy cows. Endocrinology 153:5834–5844. doi:10.1210/en.2012-1765.

Gillund, P., O. Reksen, Y.T. Gröhn, and K. Karlberg. 2001. Body condition related to ketosis and reproductive performance in Norwegian dairy cows. J. Dairy Sci.

84:1390–1396. doi:10.3168/jds.S0022-0302(01)70170-1.

Gloaguen, Y., F. Morton, R. Daly, R. Gurden, S. Rogers, J. Wandy, D. Wilson, M. Barrett, and K. Burgess. 2017. PiMP my metabolome: An integrated, web-based tool for LC-MS metabolomics data. Bioinformatics 33:4007–4009. doi:10.1093/bioinformatics/btx499.

Gobikrushanth, M., K. Macmillan, A. Behrouzi, B. Hoff, and M.G. Colazo. 2019. The factors associated with postpartum body condition score change and its relationship with serum analytes, milk production and reproductive performance in dairy cows. Livest. Sci. 228:151–160. doi:10.1016/j.livsci.2019.05.016.

Golbraikh, A., and A. Tropsha. 2002. Beware of q2!. J. Mol. Graph. Model. 20:269–276. doi:10.1016/s1093-3263(01)00123-1.

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

Görg, A., W. Weiss, and M.J. Dunn. 2004. Current two-dimensional electrophoresis technology for proteomics. Proteomics 4:3665–3685. doi:10.1002/pmic.200401031.

Goshima, N., Y. Kawamura, A. Fukumoto, A. Miura, R. Honma, R. Satoh, A. Wakamatsu, J.I. Yamamoto, K. Kimura, T. Nishikawa, T. Andoh, Y. Iida, K. Ishikawa, E. Ito, N. Kagawa, C. Kaminaga, K.I. Kanehori, B. Kawakami, K. Kenmochi, R. Kimura, M. Kobayashi, T. Kuroita, H. Kuwayama, Y. Maruyama, K. Matsuo, K. Minami, M. Mitsubori, M. Mori, R. Morishita, A. Murase, A. Nishikawa, S. Nishikawa, T. Okamoto, N. Sakagami, Y. Sakamoto, Y. Sasaki, T. Seki, S. Sono, A. Sugiyama, T. Sumiya, T. Takayama, Y. Takayama, H. Takeda, T. Togashi, K. Yahata, H. Yamada, Y. Yanagisawa, Y. Endo, F. Imamoto, Y. Kisu, S. Tanaka, T. Isogai, J.I. Imai, S. Watanabe, and N. Nomura. 2008. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat. Methods 5:1011–1017. doi:10.1038/nmeth.1273.

Grummer, R.R. 1993. Etiology of lipid-related metabolic disorders in periparturient dairy cows. J. Dairy Sci. 76:3882–3896. doi:10.3168/jds.S0022-0302(93)77729-2.

Grummer, R.R., and R.R. Rastani. 2004. Why reevaluate dry period length?. J. Dairy Sci. 87:E77–E85. doi:10.3168/jds.S0022-0302(04)70063-6.

Gu, Z., R. Eils, and M. Schlesner. 2016. Complex heatmaps reveal patterns and

correlations in multidimensional genomic data.. Bioinformatics 32:2847–2849. doi:10.1093/bioinformatics/btw313.

Han, S.J., A. Glatman Zaretsky, V. Andrade-Oliveira, N. Collins, A. Dzutsev, J. Shaik, D. Morais da Fonseca, O.J. Harrison, S. Tamoutounour, A.L. Byrd, M. Smelkinson, N. Bouladoux, J.B. Bliska, J.M. Brenchley, I.E. Brodsky, and Y. Belkaid. 2017. White Adipose Tissue Is a Reservoir for Memory T Cells and Promotes Protective Memory Responses to Infection. Immunity 47:1154-1168.e6. doi:10.1016/j.immuni.2017.11.009.

Han, X., A. Aslanian, and J.R. Yates. 2008. Mass spectrometry for proteomics. Curr. Opin. Chem. Biol. 12:483–490. doi:10.1016/j.cbpa.2008.07.024.

Harrigan, G.G., and R. Goodacre. 2003. Metabolic profiling: its role in biomarker discovery and gene function analysis. Page in Springer US.

Hartwell, J.R., M.J. Cecava, and S.S. Donkin. 2000. Impact of dietary rumen undegradable protein and rumen-protected choline on intake, peripartum liver triacylglyceride, plasma metabolites and milk production in transition dairy cows. J. Dairy Sci. 83:2907–2917. doi:10.3168/jds.S0022-0302(00)75191-5.

Hastie, T., R. Tibshirani, G. Sherlock, M. Eisen, P. Brown, and D. Botstein. 2006. Imputing missing data for gene expression arrays. Technical Report, Division of Biostatistics, Stanford University.

Hirvonen, J., G. Huszenicza, M. Kulcsàr, and S. Pyörälä. 1999. Acute-phase response in dairy cows with acute postpartum metritis. Theriogenology 51:1071–1083. doi:https://doi.org/10.1016/S0093-691X(99)80012-0.

Hiss, S., C. Weinkauf, S. Hachenberg, and H. Sauerwein. 2009. Relationship between metabolic status and the milk concentrations of haptoglobin and lactoferrin in dairy cows during early lactation. J. Dairy Sci. 92:4439–4443. doi:10.3168/jds.2008-1632.

Holtenius, K., S. Agenäs, C. Delavaud, and Y. Chilliard. 2003. Effects of feeding intensity during the dry period. 2. Metabolic and hormonal responses. J. Dairy Sci. 86:883–891. doi:10.3168/jds.S0022-0302(03)73671-6.

Huang, Y., and H. Zhu. 2017. Protein array-based approaches for biomarker discovery in cancer. Genomics, Proteomics Bioinforma. 15:73–81. doi:10.1016/j.gpb.2017.03.001.

Huber, W., A. Von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron. 2002.

Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18. doi:10.1093/bioinformatics/18.suppl_1.S96.

Huzzey, J.M., T.F. Duffield, S.J. LeBlanc, D.M. Veira, D.M. Weary, and M.A.G. Von Keyserlingk. 2009. Short communication: Haptoglobin as an early indicator of metritis. J. Dairy Sci. 92:621–625. doi:10.3168/jds.2008-1526.

Huzzey, J.M., and T.R. Overton. 2013. Using physiological markers to detect health and production problems in transition dairy cows. WCDS Adv. Dairy Technol. 25:329–339.

Ingvartsen, K.L. 2006. Feeding- and management-related diseases in the transition cow: Physiological adaptations around calving and strategies to reduce feeding-related diseases. Anim. Feed Sci. Technol. 126:175–213. doi:10.1016/j.anifeedsci.2005.08.003.

Jollife, I.T., and J. Cadima. 2016. Principal component analysis: A review and recent developments. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374. doi:10.1098/rsta.2015.0202.

Kadowaki, T., and T. Yamauchi. 2005. Adiponectin and adiponectin receptors. Endocr. Rev. 26:439–451. doi:10.1210/er.2005-0005.

Kasimanickam, R., T.F. Duffield, R.A. Foster, C.J. Gartley, K.E. Leslie, J.S. Walton, and W.H. Johnson. 2004. Endometrial cytology and ultrasonography for the detection of subclinical endometritis in postpartum dairy cows. Theriogenology 62:9–23. doi:10.1016/j.theriogenology.2003.03.001.

Katajamaa, M., and M. Orešič. 2005. Processing methods for differential analysis of LC/MS profile data. BMC Bioinformatics 6. doi:10.1186/1471-2105-6-179.

Kaufman, L., and P. J. Rousseeuw.1990. Finding groups in data. John Wiley and Sons, Inc. https://doi.org/10.1002/9780470316801

Kim, J.E., and H.G. Lee. 2021. Amino acids supplementation for the milk and milk protein production of dairy cows. Animals 11. doi:10.3390/ani11072118.

Kim, Y., H. Atalla, B. Mallard, C. Robert, and N. Karrow. 2011. Changes in Holstein cow milk and serum proteins during intramammary infection with three different strains of Staphylococcus aureus. BMC Vet. Res. 7. doi:10.1186/1746-6148-7-51.

Klose, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik 26:231–243. doi:10.1007/BF00281458.

Kok, A., J. Chen, B. Kemp, and A.T.M. Van Knegsel. 2019. Review: Dry period length in dairy cows and consequences for metabolism and welfare and customised management strategies. Animal 13:S42–S51. doi:10.1017/S1751731119001174.

Kok, A., R.J. van Hoeij, B.J. Tolkamp, M.J. Haskell, A.T.M. van Knegsel, I.J.M. de Boer, and E.A.M. Bokkers. 2017. Behavioural adaptation to a short or no dry period with associated management in dairy cows. Appl. Anim. Behav. Sci. 186:7–15. doi:10.1016/j.applanim.2016.10.017.

Kokkonen, T., J. Taponen, T. Anttila, L. Syrjälä-Qvist, C. Delavaud, Y. Chilliard, M. Tuori, and A.T. Tesfa. 2005. Effect of body fatness and glucogenic supplement on lipid and protein mobilization and plasma leptin in dairy cows. J. Dairy Sci. 88:1127–1141. doi:10.3168/jds.S0022-0302(05)72779-X.

De Koster, J., C. Urh, M. Hostens, W. Van den Broeck, H. Sauerwein, and G. Opsomer. 2017. Relationship between serum adiponectin concentration, body condition score, and peripheral tissue insulin response of dairy cows during the dry period. Domest. Anim. Endocrinol. 59:100–104. doi:10.1016/j.domaniend.2016.12.004.

Kotsiantis, S.B., I.D. Zaharakis, and P.E. Pintelas. 2006. Machine learning: A review of classification and combining techniques. Artif. Intell. Rev. 26:159–190. doi:10.1007/s10462-007-9052-3.

Krause, K.M., and G.R. Oetzel. 2006. Understanding and preventing subacute ruminal acidosis in dairy herds: A review. Anim. Feed Sci. Technol. 126:215–236. doi:10.1016/j.anifeedsci.2005.08.004.

LaCount, D.W., J.K. Drackley, and D.J. Weigel. 1995. Responses of dairy cows during early lactation to ruminal or abomasal administration of L-Carnitine. J. Dairy Sci. 78:1824–1836. doi:10.3168/jds.S0022-0302(95)76807-2.

Laemmli, UK. Nature 1970; 227: 680-685.

Larsen, M., H. Lapierre, and N.B. Kristensen. 2014. Abomasal protein infusion in postpartum transition dairy cows: Effect on performance and mammary metabolism. J. Dairy Sci. 97:5608–5622. doi:10.3168/jds.2013-7247.

Law, R.H.P., Q. Zhang, S. McGowan, A.M. Buckle, G.A. Silverman, W. Wong, C.J. Rosado, C.G. Langendorf, R.N. Pike, P.I. Bird, and J.C. Whisstock. 2006. An overview of the serpin superfamily. Genome Biol. 7:1–11. doi:10.1186/gb-2006-7-5-216.

Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: An R Package for Multivariate Analysis. J. Stat. Softw. 25:1–18. doi:10.18637/jss.v025.i01.

Lee, C., A.N. Hristov, K.S. Heyler, T.W. Cassidy, H. Lapierre, G.A. Varga, and C. Parys. 2012. Effects of metabolizable protein supply and amino acid supplementation on nitrogen utilization, milk production, and ammonia emissions from manure in dairy cows. J. Dairy Sci. 95:5253–5268. doi:10.3168/jds.2012-5366.

Leon, I.R., V. Schwammle, O.N. Jensen, and R.R. Sprenger. 2013. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. Mol. Cell. Proteomics 12:2992–3005. doi:10.1074/mcp.M112.025585.

Li, K.W., M.A., Gonzalez-Lozano, F. Koopmans, and A.B. Smit. 2020. Recent Developments in Data Independent Acquisition (DIA) Mass Spectrometry: Application of Quantitative Analysis of the Brain Proteome. Front. Mol. Neurosc. 13:564446. DOI=10.3389/fnmol.2020.564446.

Li, Y., H. Ding, X. Wang, L. Liu, D. Huang, R. Zhang, L. Guo, Z. Wang, X. Li, G. Liu, J. Wu, and X. Li. 2016. High levels of acetoacetate and glucose increase expression of cytokines in bovine hepatocytes, through activation of the NF-κB signalling pathway. J. Dairy Res. 83:51–57. doi:10.1017/S0022029915000680.

Liu, Q., D. Walker, K. Uppal, Z. Liu, C. Ma, V.L. Tran, S. Li, D.P. Jones, and T. Yu. 2020. Addressing the batch effect issue for LC/MS metabolomics data in data preprocessing. Sci. Rep. 10:1–13. doi:10.1038/s41598-020-70850-0.

Liumbruno, G., A. D'Alessandro, G. Grazzini, and L. Zolla. 2010. Blood-related proteomics. J. Proteomics 73:483–507. doi:10.1016/j.jprot.2009.06.010.

Luo, Z.Z., L.H. Shen, J. Jiang, Y.X. Huang, L.P. Bai, S.M. Yu, X.P. Yao, Z.H. Ren, Y.X. Yang, and S.Z. Cao. 2019. Plasma metabolite changes in dairy cows during parturition identified using untargeted metabolomics. J. Dairy Sci. 102:4639–4650. doi:10.3168/jds.2018-15601.

Madoz, L.V., M.J. Giuliodori, M. Jaureguiberry, J. Plöntzke, M. Drillich, and R.L. de la Sota. 2013. The relationship between endometrial cytology during estrous cycle and

cutoff points for the diagnosis of subclinical endometritis in grazing dairy cows. J. Dairy Sci. 96:4333–4339. doi:10.3168/jds.2012-6269.

Mann, M., R.C. Hendrickson, and A. Pandey. 2001. Analysis of proteins and proteomes by mass spectometry. Annual Review of Biochem. Vol. 70:437-473. https://doi.org/10.1146/annurev.biochem.70.1.437.

Manza, L.L., S.L. Stamer, A.-J.L. Ham, S.G. Codreanu, and D.C. Liebler. 2005. Sample preparation and digestion for proteomic analyses using spin filters. Proteomics 5:1742–1745. doi:10.1002/pmic.200401063.

Martínez, Y., X. Li, G. Liu, P. Bin, W. Yan, D. Más, M. Valdivié, C.A.A. Hu, W. Ren, and Y. Yin. 2017. The role of methionine on metabolism, oxidative stress, and diseases. Amino Acids 49:2091–2098. doi:10.1007/s00726-017-2494-2.

Masuda, T., M. Tomita, and Y. Ishihama. 2008. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. J. Proteome Res. 7:731–740. doi:10.1021/pr700658q.

McFadden, J.W. 2020. Review: Lipid biology in the periparturient dairy cow: Contemporary perspectives. Animal 14:S165–S175. doi:10.1017/S1751731119003185.

McGuffey, R.K. 2017. A 100-Year Review: Metabolic modifiers in dairy cattle nutrition. J. Dairy Sci. 100:10113–10142. doi:10.3168/jds.2017-12987.

Meikle, A., M. Kulcsar, Y. Chilliard, H. Febel, C. Delavaud, D. Cavestany, and P. Chilibroste. 2004. Effects of parity and body condition at parturition on endocrine and reproductive parameters of the cow. Reproduction 127:727–737. doi:10.1530/rep.1.00080.

Melcher, Y., I. Prunner, and M. Drillich. 2014. Degree of variation and reproducibility of different methods for the diagnosis of subclinical endometritis. Theriogenology 82:57–63. doi:10.1016/j.theriogenology.2014.03.003.

Meyer, J., S.U. Daniels, S. Grindler, J. Tröscher-Mußotter, M. Alaedin, J. Frahm, L. Hüther, J. Kluess, S. Kersten, D. von Soosten, U. Meyer, E. Most, K. Eder, H. Sauerwein, J. Seifert, K. Huber, J. Rehage, and S. Dänicke. 2020. Effects of a dietary l-carnitine supplementation on performance, energy metabolism and recovery from calving in dairy cows. Animals 10. doi:10.3390/ani10020342.

Mielenz, M., B. Mielenz, S.P. Singh, C. Kopp, J. Heinz, S. Häussler, and H. Sauerwein. 2013. Development, validation, and pilot application of a semiquantitative Western blot analysis and an ELISA for bovine adiponectin. Domest. Anim. Endocrinol. 44:121–130. doi:10.1016/j.domaniend.2012.10.004.

Miller, B.A., A. Brewer, P. Nanni, J.J. Lim, J.J. Callanan, J. Grossmann, L. Kunz, A.M. de Almeida, K.G. Meade, and A. Chapwanya. 2019. Characterization of circulating plasma proteins in dairy cows with cytological endometritis. J. Proteomics 205:103421. doi:10.1016/j.jprot.2019.103421.

Millioni, R., S. Tolin, L. Puricelli, S. Sbrignadello, G.P. Fadini, P. Tessari, and G. Arrigoni. 2011. High abundance proteins depletion vs low abundance proteins enrichment: Comparison of methods to reduce the plasma proteome complexity. PLoS One 6. doi:10.1371/journal.pone.0019603.

Min, L., J. Cheng, S. Zhao, H. Tian, Y. Zhang, S. Li, H. Yang, N. Zheng, and J. Wang. 2016. Plasma-based proteomics reveals immune response, complement and coagulation cascades pathway shifts in heat-stressed lactating dairy cows. J. Proteomics 146:99–108. doi:10.1016/j.jprot.2016.06.008.

Morrison, C.D., J.A. Daniel, B.J. Holmberg, J. Djiane, N. Raver, A. Gertler, and D.H. Keilser. 2001. Central infusion of leptin into well-fed and undernourished ewe lambs: Effects on feed intake and serum concentrations of growth hormone and luteinizing hormone. J. Endocrinol. 168:317–324. doi:10.1677/joe.0.1680317.

Ng, K.K.S., K. Drickamer, and W.I. Weis. 1996. Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. J. Biol. Chem. 271:663–674. doi:10.1074/jbc.271.2.663.

Nowroozi-Asl, A., S. Nazifi, A. R. Ghasrodashti, and A. Olyaee. 2011. Prevalence of subclinical ketosis in dairy cattle in the Southwestern Iran and detection of cutoff point for NEFA and glucose concentrations for diagnosis of subclinical ketosis. Prev. Vet. Med. 100(1):38-43. https://doi.org/10.1016/j.prevetmed.2011.02.013

Nijdam, A.J., M.R. Zianni, E.E. Herderick, M.M.C. Cheng, J.R. Prosperi, F.A. Robertson, E.F. Petricoin, L.A. Liotta, and M. Ferrari. 2009. Application of physicochemically modified silicon substrates as reverse-phase protein microarrays. J. Proteome Res. 8:1247–1254. doi:10.1021/pr800455y.

Oliver, S.G., M.K. Winson, D.B. Kell, and F. Baganz. 1998. Systematic functional analysis

of the yeast genome. Trends Biotechnol. 16:373–378. doi:10.1016/S0167-7799(98)01214-1.

Osorio, J.S., P. Ji, J.K. Drackley, D. Luchini, and J.J. Loor. 2013. Erratum to "Supplemental Smartamine M or MetaSmart during the transition period benefits postpartal cow performance and blood neutrophil function" (J. Dairy Sci. 96:6248–6263). J. Dairy Sci. 96:8093. doi:10.3168/jds.2013-96-12-8093.

Overton, T.R., and M.R. Waldron. 2004. Nutritional management of transition dairy cows: Strategies to optimize metabolic health. J. Dairy Sci. 87:E105–E119. doi:10.3168/jds.S0022-0302(04)70066-1.

Package, T., and A.C. Lazar. 2015. Package ' imputeLCMD ' 1–30.

Pappritz, J., U. Meyer, R. Kramer, E.M. Weber, G. Jahreis, J. Rehage, G. Flachowsky, and S. Dänicke. 2011. Effects of long-term supplementation of dairy cow diets with rumenprotected conjugated linoleic acids (CLA) on performance, metabolic parameters and fatty acid profile in milk fat. Arch. Anim. Nutr. 65:89–107. doi:10.1080/1745039X.2011.552275.

Phizicky, E., P.I.H. Bastiaens, H. Zhu, M. Snyder, and S. Fields. 2003. Protein analysis on a proteomic scale. Nature 422:208–215. doi:10.1038/nature01512.

Piepenbrink, M.S., A.L. Marr, M.R. Waldron, W.R. Butler, T.R. Overton, M. Vázquez-Añón, and M.D. Holt. 2004. Feeding 2-Hydroxy-4-(Methylthio)-Butanoic acid to periparturient dairy cows improves milk production but not hepatic metabolism. J. Dairy Sci. 87:1071–1084. doi:10.3168/jds.S0022-0302(04)73253-1.

Pino, L. K., S.C. Just, M.J., MacCoss, and B.C. Searle. 2020. Acquiring and analyzing data independent acquisition proteomics experiments without spectrum libraries. Mol. Cell Proteomics 19, 1088–1103. doi: 10.1074/mcp.P119.001913.

Pires, J.A.A., C. Delavaud, Y. Faulconnier, D. Pomiès, and Y. Chilliard. 2013. Effects of body condition score at calving on indicators of fat and protein mobilization of periparturient Holstein-Friesian cows. J. Dairy Sci. 96:6423–6439. doi:10.3168/jds.2013-6801.

Pirestani, A., and M. Aghakhani. 2018. The effects of rumen-protected choline and L-carnitine supplementation in the transition period on reproduction, production, and some metabolic diseases of dairy cattle. J. Appl. Anim. Res. 46:435–440. doi:10.1080/09712119.2017.1332632.

Pohl, A., O. Burfeind, and W. Heuwieser. 2015. The associations between postpartum serum haptoglobin concentration and metabolic status, calving difficulties, retained fetal membranes, and metritis. J. Dairy Sci. 98:4544–4551. doi:10.3168/jds.2014-9181.

Preynat, A., H. Lapierre, M.C. Thivierge, M.F. Palin, J.J. Matte, A. Desrochers, and C.L. Girard. 2009. Effects of supplements of folic acid, vitamin B12, and rumen-protected methionine on whole body metabolism of methionine and glucose in lactating dairy cows. J. Dairy Sci. 92:677–689. doi:10.3168/jds.2008-1525.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Accessed June. 10, 2019. https://www.R-project.org.

Rathbun, F.M., R.S. Pralle, S.J. Bertics, L.E. Armentano, K. Cho, C. Do, K.A. Weigel, and H.M. White. 2017. Relationships between body condition score change, prior mid-lactation phenotypic residual feed intake, and hyperketonemia onset in transition dairy cows. J. Dairy Sci. 100:3685–3696. doi:10.3168/jds.2016-12085.

Reist, M., D. Erdin, D. Von Euw, K. Tschuemperlin, H. Leuenberger, C. Delavaud, Y. Chilliard, H.M. Hammon, N. Kuenzi, and J.W. Blum. 2003. Concentrate feeding strategy in lactating dairy cows: Metabolic and endocrine changes with emphasis on leptin. J. Dairy Sci. 86:1690–1706. doi:10.3168/jds.S0022-0302(03)73755-2.

Reynolds, C.K., P.C. Aikman, B. Lupoli, D.J. Humphries, and D.E. Beever. 2003. Splanchnic metabolism of dairy cows during the transition from late gestation through early lactation. J. Dairy Sci. 86:1201–1217. doi:10.3168/jds.S0022-0302(03)73704-7.

Ritchie, M.E., B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, and G.K. Smyth. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43:e47. doi:10.1093/nar/gkv007.

Roberts, L.D., A.L. Souza, R.E. Gerszten, and C.B. Clish. 2012. Targeted metabolomics. Curr Protoc Mol Biol. Chapter 30:Unit 30.2.1-24. doi: 10.1002/0471142727.mb3002s98.

Roche, J.R., D.P. Berry, J.M. Lee, K.A. Macdonald, and R.C. Boston. 2007a. Describing the body condition score change between successive calvings: A novel strategy generalizable to diverse cohorts. J. Dairy Sci. 90:4378–4396. doi:10.3168/jds.2006-729.

Roche, J.R., N.C. Friggens, J.K. Kay, M.W. Fisher, K.J. Stafford, and D.P. Berry. 2009. Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. J. Dairy Sci. 92:5769–5801. doi:10.3168/jds.2009-2431.

Roche, J.R., J.K. Kay, A.G. Rius, T.M. Grala, A.J. Sheahan, H.M. White, and C.V.C. Phyn. 2013a. Short communication: Immediate and deferred milk production responses to concentrate supplements in cows grazing fresh pasture. J. Dairy Sci. 96:2544–2550. doi:10.3168/jds.2012-4626.

Roche, J.R., J.M. Lee, K.A. Macdonald, and D.P. Berry. 2007b. Relationships among body condition score, body weight, and milk production variables in pasture-based dairy cows. J. Dairy Sci. 90:3802–3815. doi:10.3168/jds.2006-740.

Roche, J.R., K.A. Macdonald, C.R. Burke, J.M. Lee, and D.P. Berry. 2007c. Associations among body condition score, body weight, and reproductive performance in seasonal-calving dairy cattle. J. Dairy Sci. 90:376–391. doi:10.3168/jds.s0022-0302(07)72639-5.

Roche, J.R., K.A. Macdonald, K.E. Schütz, L.R. Matthews, G.A. Verkerk, S. Meier, J.J. Loor, A.R. Rogers, J. McGowan, S.R. Morgan, S. Taukiri, and J.R. Webster. 2013b. Calving body condition score affects indicators of health in grazing dairy cows. J. Dairy Sci. 96:5811–5825. doi:10.3168/jds.2013-6600.

Roncada, P., C. Piras, A. Soggiu, R. Turk, A. Urbani, and L. Bonizzi. 2012. Farm animal milk proteomics. J. Proteomics 75:4259–4274. doi:10.1016/j.jprot.2012.05.028.

Rousseeuw, P.J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20:53–65. doi:10.1016/0377-0427(87)90125-7.

Ruiz-Perez, D., H. Guan, P. Madhivanan, K. Mathee, and G. Narasimhan. 2020. So you think you can PLS-DA?. BMC Bioinformatics 21:1–10. doi:10.1186/s12859-019-3310-7.

Saremi, B., A. Al-Dawood, S. Winand, U. Müller, J. Pappritz, D. von Soosten, J. Rehage, S. Dänicke, S. Häussler, M. Mielenz, and H. Sauerwein. 2012. Bovine haptoglobin as an adipokine: Serum concentrations and tissue expression in dairy cows receiving a conjugated linoleic acids supplement throughout lactation. Vet. Immunol. Immunopathol. 146:201–211. doi:10.1016/j.vetimm.2012.03.011.

Sauerwein, H., and S. Häußler. 2016. Endogenous and exogenous factors influencing the

concentrations of adiponectin in body fluids and tissues in the bovine. Domest. Anim. Endocrinol. 56:S33–S43. doi:10.1016/j.domaniend.2015.11.007.

Sauerwein, H., U. Heintges, M. Hennies, T. Selhorst, and A. Daxenberger. 2004. Growth hormone induced alterations of leptin serum concentrations in dairy cows as measured by a novel enzyme immunoassay. Livest. Prod. Sci. 87:189–195. doi:10.1016/j.livprodsci.2003.08.001.

Scalbert, A., L. Brennan, O. Fiehn, T. Hankemeier, B.S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart, and S. Wopereis. 2009. Mass-spectrometry-based metabolomics: Limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics 5:435–458. doi:10.1007/s11306-009-0168-0.

Scheltema, R.A., A. Jankevics, R.C. Jansen, M.A. Swertz, and R. Breitling. 2011. PeakML/mzMatch: A file format, Java library, R library, and tool-chain for mass spectrometry data analysis. Anal. Chem. 83:2786–2793. doi:10.1021/ac2000994.

Schrimpe-Rutledge, A.C., S.G. Codreanu, S.D. Sherrod, and J.A. McLean. 2016. Untargeted metabolomics strategies-challenges and emerging directions.. J. Am. Soc. Mass Spectrom. 27:1897–1905. doi:10.1007/s13361-016-1469-y.

Schröder, U.J., and R. Staufenbiel. 2006. Invited review: Methods to determine body fat reserves in the dairy cow with special regard to ultrasonographic measurement of backfat thickness. J. Dairy Sci. 89:1–14. doi:10.3168/jds.S0022-0302(06)72064-1.

Schuh, K., H. Sadri, S. Häussler, L.A. Webb, C. Urh, M. Wagner, C. Koch, J. Frahm, S. Dänicke, G. Dusel, and H. Sauerwein. 2019. Comparison of performance and metabolism from late pregnancy to early lactation in dairy cows with elevated v. normal body condition at dry-off. Animal 13:1478–1488. doi:10.1017/S1751731118003385.

Schwab, C.G., and G.A. Broderick. 2017. A 100-Year Review: Protein and amino acid nutrition in dairy cows. J. Dairy Sci. 100:10094–10112. doi:10.3168/jds.2017-13320.

Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. J Am Soc Mass Spectrom. 2016 Dec;27(12):1897-1905. doi: 10.1007/s13361-016-1469-y.

Shin, D.H., J.K. Jeong, I.S. Choi, S.H. Moon, S.C. Lee, H.G. Kang, S.B. Park, and I.H. Kim. 2018. Associations between serum haptoglobin concentration and peri- and

postpartum disorders, milk yield, and reproductive performance in dairy cows. Livest. Sci. 213:14–18. doi:10.1016/j.livsci.2018.04.015.

Singh, B., G. Mal, S.K. Gautam, and M. Mukesh. 2019. Advances in Animal Biotechnology.

Singh, S.P., S. Häussler, J.F.L. Heinz, B. Saremi, B. Mielenz, J. Rehage, S. Dänicke, M. Mielenz, and H. Sauerwein. 2014. Supplementation with conjugated linoleic acids extends the adiponectin deficit during early lactation in dairy cows. Gen. Comp. Endocrinol. 198:13–21. doi:10.1016/j.ygcen.2013.12.008.

Smith, C.A., E.J. Want, G. O'Maille, R. Abagyan, and G. Siuzdak. 2006. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. 78:779–787. doi:10.1021/ac051437y.

Smith, T.H., and J.P. McNamara. 1990. Regulation of bovine adipose tissue metabolism during lactation. 6. Cellularity and hormone-sensitive lipase activity as affected by genetic merit and energy intake. J. Dairy Sci. 73:772–783. doi:10.3168/jds.S0022-0302(90)78730-9.

Soares, R., C. Franco, E. Pires, M. Ventosa, R. Palhinhas, K. Koci, A. Martinho de Almeida, and A. Varela Coelho. 2012. Mass spectrometry and animal science: Protein identification strategies and particularities of farm animal species. J. Proteomics 75:4190–4206. doi:10.1016/j.jprot.2012.04.009.

Sordillo, L.M., G.A. Contreras, and S.L. Aitken. 2009. Metabolic factors affecting the inflammatory response of periparturient dairy cows.. Anim. Health Res. Rev. 10:53–63. doi:10.1017/S1466252309990016.

Sotelo, J., and C.M. Slupsky. 2013. Metabolomics using nuclear magnetic resonance (NMR). Elsevier Masson SAS.

Süss, D., M. Iwersen, V. Schweinzer, E. Gusterer, P. Kanz, S. Krieger, H. Pothmann, K. Wagener, M. Hoelker, D. Tesfaye, K. Schellander, A. Helmbrecht, C. Parys, and M. Drillich. 2019. Supplementing rumen-protected methionine to lactating multiparous dairy cows did not improve reproductive performance. Reprod. Domest. Anim. 54:1265–1273. doi:10.1111/rda.13509.

Sutandy, F.X.R., J. Qian, C.S. Chen, and H. Zhu. 2013. Overview of protein microarrays. Curr. Protoc. Protein Sci. 1–16. doi:10.1002/0471140864.ps2701s72.

Swartz, T.H., U. Moallem, H. Kamer, G. Kra, Y. Levin, L.K. Mamedova, B.J. Bradford, and M. Zachut. 2021. Characterization of the liver proteome in dairy cows experiencing negative energy balance at early lactation. J. Proteomics 246:104308. doi:10.1016/j.jprot.2021.104308.

Takiya, C.S., S.R. Montgomery, L.K. Mamedova, G. Kra, N. Nemes-Navon, Y. Levin, S.D. Fleming, B.J. Bradford, and M. Zachut. 2019. Proteomic analysis reveals greater abundance of complement and inflammatory proteins in subcutaneous adipose tissue from postpartum cows treated with sodium salicylate. J. Proteomics 204:103399. doi:10.1016/j.jprot.2019.103399.

Torres, E., M. Mellado, C. Leyva, J.E. García, F.G. Véliz, and J. Hernández-Bustamante. 2020. Serum metabolites and body condition score associated with metritis, endometritis, ketosis, and mastitis in Holstein cows. Pesqui. Agropecu. Bras. 55. doi:10.1590/S1678-3921.PAB2020.V55.01308.

Tu, C., P.A. Rudnick, M.Y. Martinez, K.L. Cheek, S.E. Stein, R.J.C. Slebos, and D.C. Liebler. 2011. Depletion of abundant plasma proteins and limitations of plasma proteomics. J. Proteom. Research 9:4982–4991. doi:10.1021/pr100646w.Depletion.

Tukey, J.W. 1977. Exploratory Data Analysis by John W. Tukey. Biometrics 33:768.

Ünlü, M., M.E. Morgan, and J.S. Minden. 1997. Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. Electrophoresis 18:2071–2077. doi:10.1002/elps.1150181133.

Vailati-Riboni, M., Z. Zhou, C.B. Jacometo, A. Minuti, E. Trevisi, D.N. Luchini, and J.J. Loor. 2017. Supplementation with rumen-protected methionine or choline during the transition period influences whole-blood immune response in periparturient dairy cows. J. Dairy Sci. 100:3958–3968. doi:10.3168/jds.2016-11812.

Valletta, J.J., C. Torney, M. Kings, A. Thornton, and J. Madden. 2017. Applications of machine learning in animal behaviour studies. Anim. Behav. 124:203–220. doi:10.1016/j.anbehav.2016.12.005.

De Veth, M.J., J.M. Griinari, A.M. Pfeiffer, and D.E. Bauman. 2004. Effect of CLA on milk fat synthesis in dairy cows: Comparison of inhibition by methyl esters and free fatty acids, and relationships among studies. Lipids 39:365–372. doi:10.1007/s11745-004-1240-4.

Wasinger, V.C., S.J. Cordwell, A. Poljak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W.

Duncan, R. Harris, K.L. Williams, and I. Humphery-Smith. 1995. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis 16:1090–1094. doi:10.1002/elps.11501601185.

Weber, C., C. Hametner, A. Tuchscherer, B. Losand, E. Kanitz, W. Otten, S.P. Singh, R.M. Bruckmaier, F. Becker, W. Kanitz, and H.M. Hammon. 2013. Variation in fat mobilization during early lactation differently affects feed intake, body condition, and lipid and glucose metabolism in high-yielding dairy cows. J. Dairy Sci. 96:165–180. doi:10.3168/jds.2012-5574.

Wickham H (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wilkins, M.R. 2009. Hares and tortoises: The high- versus low-throughput proteomic race. Electrophoresis 30:150–155. doi:10.1002/elps.200900175.

Wilkins, M.R., J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, and K.L. Williams. 1996. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. Rev. 13:19–50. doi:10.1080/02648725.1996.10647923.

Williams, E.J., D.P. Fischer, D.U. Pfeiffer, G.C.W. England, D.E. Noakes, H. Dobson, and I.M. Sheldon. 2005. Clinical evaluation of postpartum vaginal mucus reflects uterine bacterial infection and the immune response in cattle. Theriogenology 63:102–117. doi:10.1016/j.theriogenology.2004.03.017.

Wiśniewski, J.R., A. Zougman, N. Nagaraj, and M. Mann. 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6:359–362. doi:10.1038/nmeth.1322.

Xiao, J.F., B. Zhou, and H.W. Ressom. 2012. Metabolite identification and quantitation in LC-MS/MS-based metabolomics. Trends Analyt. Chem. 32:1–14. doi:10.1016/j.trac.2011.08.009.

Yamauchi, T., J. Kamon, H. Waki, Y. Terauchi, N. Kubota, K. Hara, Y. Mori, T. Ide, K. Murakami, N. Tsuboyama-Kasaoka, O. Ezaki, Y. Akanuma, O. Gavrilova, C. Vinson, M.L. Reitman, H. Kagechika, K. Shudo, M. Yoda, Y. Nakano, K. Tobe, R. Nagai, S. Kimura, M. Tomita, P. Froguel, and T. Kadowaki. 2001. The fat-derived hormone adiponectin reverses insulin resistance associated with both lipoatrophy and obesity. Nat. Med. 7:941–946. doi:10.1038/90984.

Yasui, T., K. McCann, R.O. Gilbert, D. V. Nydam, and T.R. Overton. 2014. Associations of cytological endometritis with energy metabolism and inflammation during the periparturient period and early lactation in dairy cows. J. Dairy Sci. 97:2763–2770. doi:10.3168/jds.2013-7322.

Zahra, L.C., T.F. Duffield, K.E. Leslie, T.R. Overton, D. Putnam, and S.J. LeBlanc. 2006. Effects of rumen-protected choline and monensin on milk production and metabolism of periparturient dairy cows. J. Dairy Sci. 89:4808–4818. doi:10.3168/jds.S0022-0302(06)72530-9.

Zebeli, Q., K. Ghareeb, E. Humer, B.U. Metzler-Zebeli, and U. Besenfelder. 2015. Nutrition, rumen health and inflammation in the transition period and their role on overall health and fertility in dairy cows. Res. Vet. Sci. 103:126–136. doi:10.1016/j.rvsc.2015.09.020.

Zebeli, Q., and B.U. Metzler-Zebeli. 2012. Interplay between rumen digestive disorders and diet-induced inflammation in dairy cattle. Res. Vet. Sci. 93:1099–1108. doi:10.1016/j.rvsc.2012.02.004.

Zhang, A., H. Sun, P. Wang, Y. Han, and X. Wang. 2012. Modern analytical techniques in metabolomics analysis. Analyst 137:293–300. doi:10.1039/C1AN15605E.

Zhou, J., H. Liu, Y. Liu, J. Liu, X. Zhao, and Y. Yin. 2016a. Development and evaluation of a parallel reaction monitoring strategy for large-scale targeted metabolomics Quantification. Anal. Chem. 88:4478–4486. doi:10.1021/acs.analchem.6b00355.

Zhou, Y., J. Ren, T. Song, J. Peng, and H. Wei. 2016b. Methionine regulates mTORC1 via the T1R1/T1R3-PLC β-Ca2++-ERK1/2 signal transduction process in C2C12 cells. Int. J. Mol. Sci. 17:1–11. doi:10.3390/ijms17101684.

Zhou, Z., O. Bulgari, M. Vailati-Riboni, E. Trevisi, M.A. Ballou, F.C. Cardoso, D.N. Luchini, and J.J. Loor. 2016c. Rumen-protected methionine compared with rumen-protected choline improves immunometabolic status in dairy cows during the peripartal period. J. Dairy Sci. 99:8956–8969. doi:10.3168/jds.2016-10986.